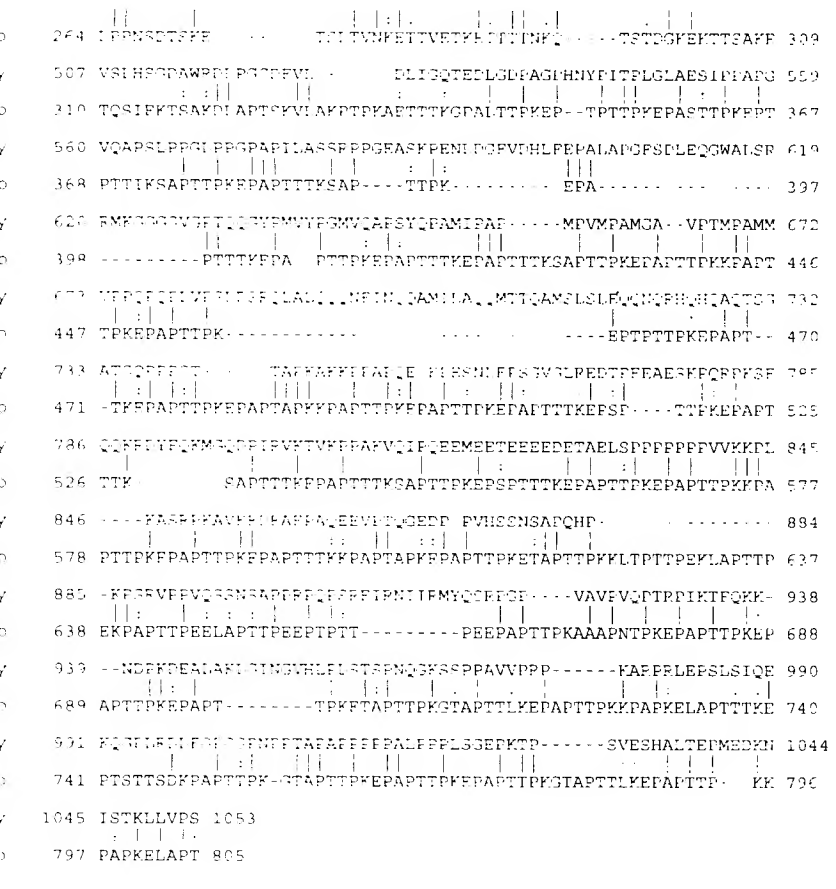
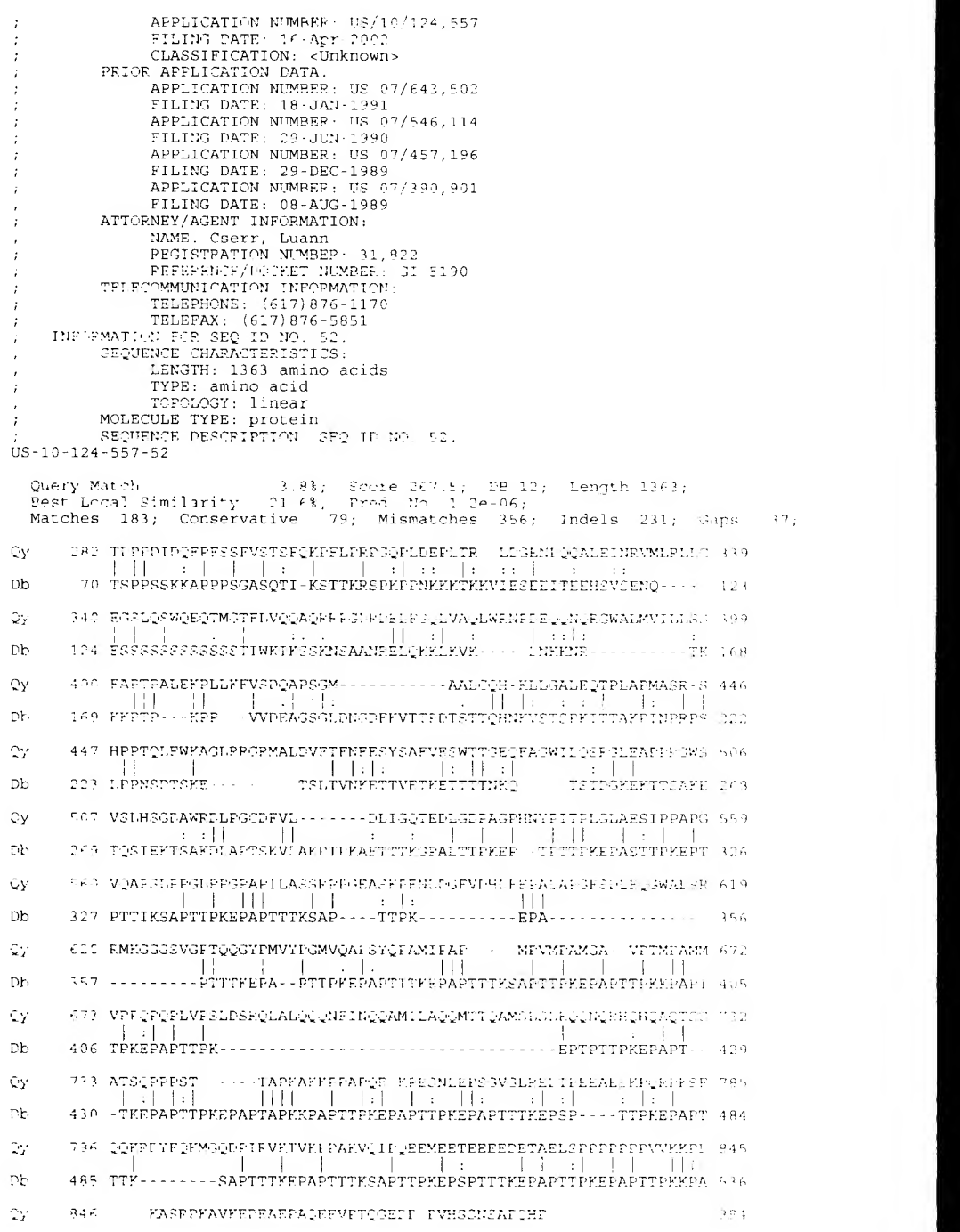


ESULT 8
S-10-124-557-52
Sequence 52, Application US/10124557
Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA

; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,922
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-10-124-557-52

Query Match 3.8%; Score 267.5; DB 12; Length 1363;
Best Local Similarity 21.6%; Pred. No. 1.2e-06;
Matches 183; Conservative 79; Mismatches 356; Indels 231; Gaps 37;



```
Db    537 PTTPKEEPAPTTKLPAPTTTKEPAPTAPKEPAPTTPKETAPTTPKELTPTTPKELAPTTP 596

Qy    880 KPSRVPPVQCCNCAPPRPQPPPEIPNIIPMYVTRRP   VAVPVSPTPPIKTPQKK  938

Db    597 EKPAPTTPEELAPTTPEEPTPTT----------PEEPAPTTPKAAAPNTPKEPAPTTPKEP 647

Qy    939 --NDPKDEALAKLGINGVHLPGNIQHNQCRGGPPAVVPEE------PAPPPLPPSGSIQP 990

Db    648 APTTPKEPAPT        TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKE 699

Qy    991 KQGPLRDLFGPCCSPNPPTAPAPPPPPALPPPLSGEPKTP       CVECHALTEPMPDYN 1044

Db    700 PTSTTSDKPAPTTPK-GTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP---KK 755

Qy   1045 ISTKLLVPS 1053

Db    756 PAPKELAPT 764
```

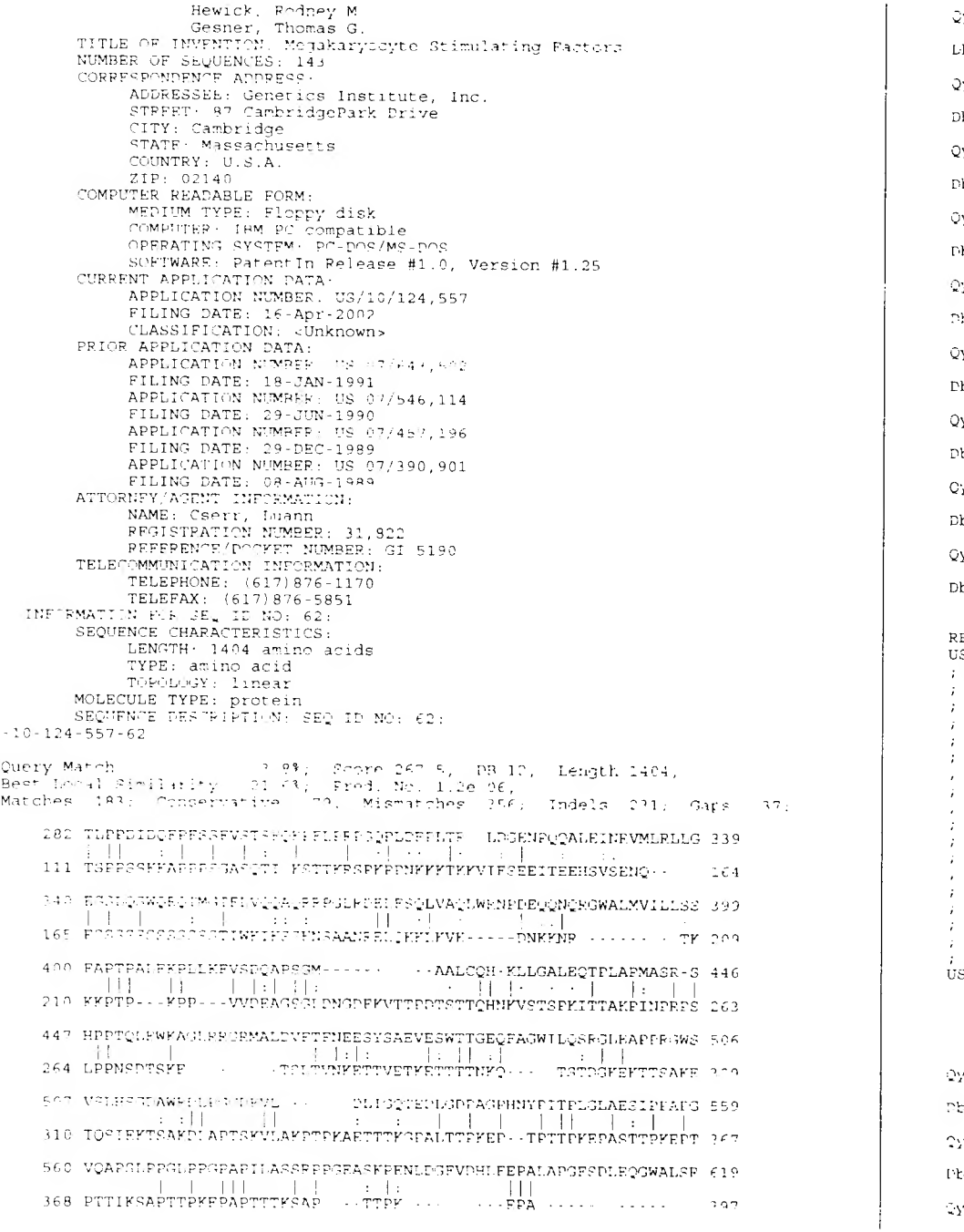
RESULT 9
US-10-124-557-2
; Sequence 2, Application US/10124557
; Patent No. US20020137894A1
;   GENERAL INFORMATION:
;   APPLICANT: Turner, Katherine
;   Clark, Stephen C.
;   Jacobs, Kenneth
;   Hewick, Rodney M.
;   Gesner, Thomas G.
;   TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;   NUMBER OF SEQUENCES: 143
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genetics Institute, Inc.
;   STREET: 87 CambridgePark Drive
;   CITY: Cambridge
;   STATE: Massachusetts
;   COUNTRY: U.S.A.
;   ZIP: 02140
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/124,557
;   FILING DATE: 16 Apr 2002
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/643,502
;   FILING DATE: 18 JAN 1991
;   APPLICATION NUMBER: US 07/546,114
;   FILING DATE: 29 JUN 1990
;   APPLICATION NUMBER: US 07/457,196
;   FILING DATE: 29 DEC 1989
;   APPLICATION NUMBER: US 07/390,901
;   FILING DATE: 08-AUG-1989
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Oserr, Luann
;   REGISTRATION NUMBER: 31,822
;   REFERENCE/DOCKET NUMBER: GI 5130
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)876-1170
;   TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1404 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-124-557-2

```
  Query Match             3.8%;  Score 267.5;  DB 1;  Length 1404;
  Best Local Similarity   21.6%;  Pred. No. 1.2e-26;
  Matches  183;  Conservative   79;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy    282 TLPPDIDQFPFSSFVSTSFQKPPLPPPQQPIDPPETR [illegible]ALEINPVMLPLL [illegible]

Db    111 TSPPSSKKAPPPSGASQII KSIIKKSPKPPNKKKIKKVIKPKKITKKPSVIENQ      164

Qy    340 EGSLQSWQEQTMGTPLVQQAQPPPGIRPPELPSQLVAQLWENI [illegible]WALMVI [illegible]

Db    165 ESSSSSSSSSSSIIWKIKSSKNSAANKK[illegible]KKIKVK      [illegible]NKKNK          [illegible]

Qy    400 EAPTPALEKPLLKPVSPQAPSCM              AA[illegible]  446

Db    210 KKPIP--KIP---VVPEA[illegible]TTAKPINPKK[illegible]

Qy    447 HPPTQLEWKAGLRRGRMALDVPTENEESYSAEVESWTT[illegible]WS [illegible]

Db    264 LPPNSDTSKE           TSLTVNKETTVETKETTTTNKQ      TSTPSKKETTIAKK [illegible]

Qy    507 VSLHSGDAWRDLPGCDPVL          DLIGQTEDI[illegible]HRNYL[illegible]GLAKS[illegible]

Db    310 TQSIKKISAKPLAPISKVIAKPIPKAKIITKPAPTTPKKP [illegible]KKPAPTTPKKPT [illegible]

Qy    560 VQAPSLPPGLPPGPAPILASSPPQGASPPEN[illegible]WALL [illegible]

Db    368 PTTIKSAPTTPKEPAPTTTKSAP    TTPK               PEA               [illegible]

Qy    620 EMKGCGSVGPTQQIYPMVYPGMVQAPCYQPAMIPKP        M[illegible]AM[illegible]MPAMM [illegible]

Db    398 ----      --PTTTKEPA   PTTPKEPAPTTTKEPAPTT[illegible]EKPAPTTPKEPAPT 440

Qy    673 VPPQPQPLVPSIPSPQLALQQQNEINQQAMILAQQMTT[illegible]

Db    447 TPKEPAPTTPK                                   [illegible]TTPKEPAPT   4[illegible]

Qy    733 ATSQPPPSI ----[illegible]APKAKKPPAPQE KPE[illegible]NLE[illegible]

Db    471 -TKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSI      TTPKEPAPT [illegible]

Qy    786 [illegible]

Db    526 TIK------ -[illegible]PAPTTTKEPAPTTTKEPAPTTTKEPA [illegible]

Qy    846      KASRPKAVKEDEAKPAQKEVPT[illegible]PVH[illegible]          [illegible]

Db    578 PTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKE[illegible]TTPKEPAPTTP [illegible]

Qy    885 -KPSRVPPVQCCNCAPPRPQPPPEIPNIIPMYQTRRPP      VAV[illegible]KTPQKK [illegible]

Db    638 EKPAPTTPEELAPTTPEEPTPTT              PEEPAPTTPKAAAPNTPKEPAPTTPKEP [illegible]

Qy    939 --NDPKDEALAKLGINGVHLPGNIQHNQCRGGPPAVVPEE         PAPPPLPPSGSIQP [illegible]

Db    689 APTTPKEPAPT    ----  TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKE [illegible]

Qy    991 KQGPLRDLFGPCCSPNPPTAPAPPPPPALPPPLSGEPKTP       CVECHALTEPMPDYN 1044

Db    741 PTSTTSDKPAPTTPK-GTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP    KK [illegible]

Qy   1045 ISTKLLVPS 1053

Db    797 PAPKELAPT 805
```

RESULT 10
US-10-124-557-62
; Sequence 62, Application US/10124557
; Patent No. US20020137894A1
;   GENERAL INFORMATION:
;   APPLICANT: Turner, Katherine
;   Clark, Stephen C.
;   Jacobs, Kenneth

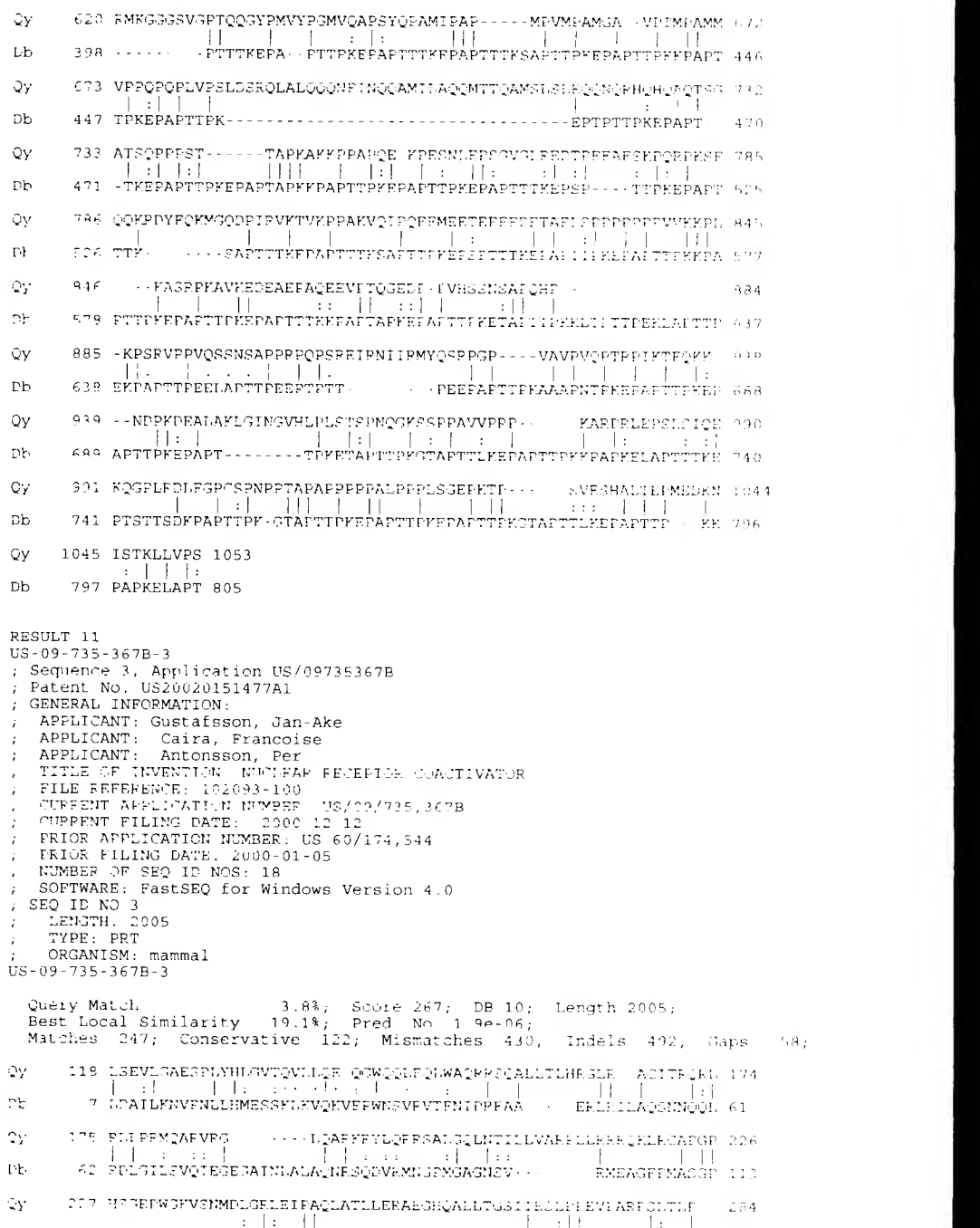
Hewick, Rodney M
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Generics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/644,602
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Csorr, Luann
REGISTRATION NUMBER: 31,922
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
-10-124-557-62

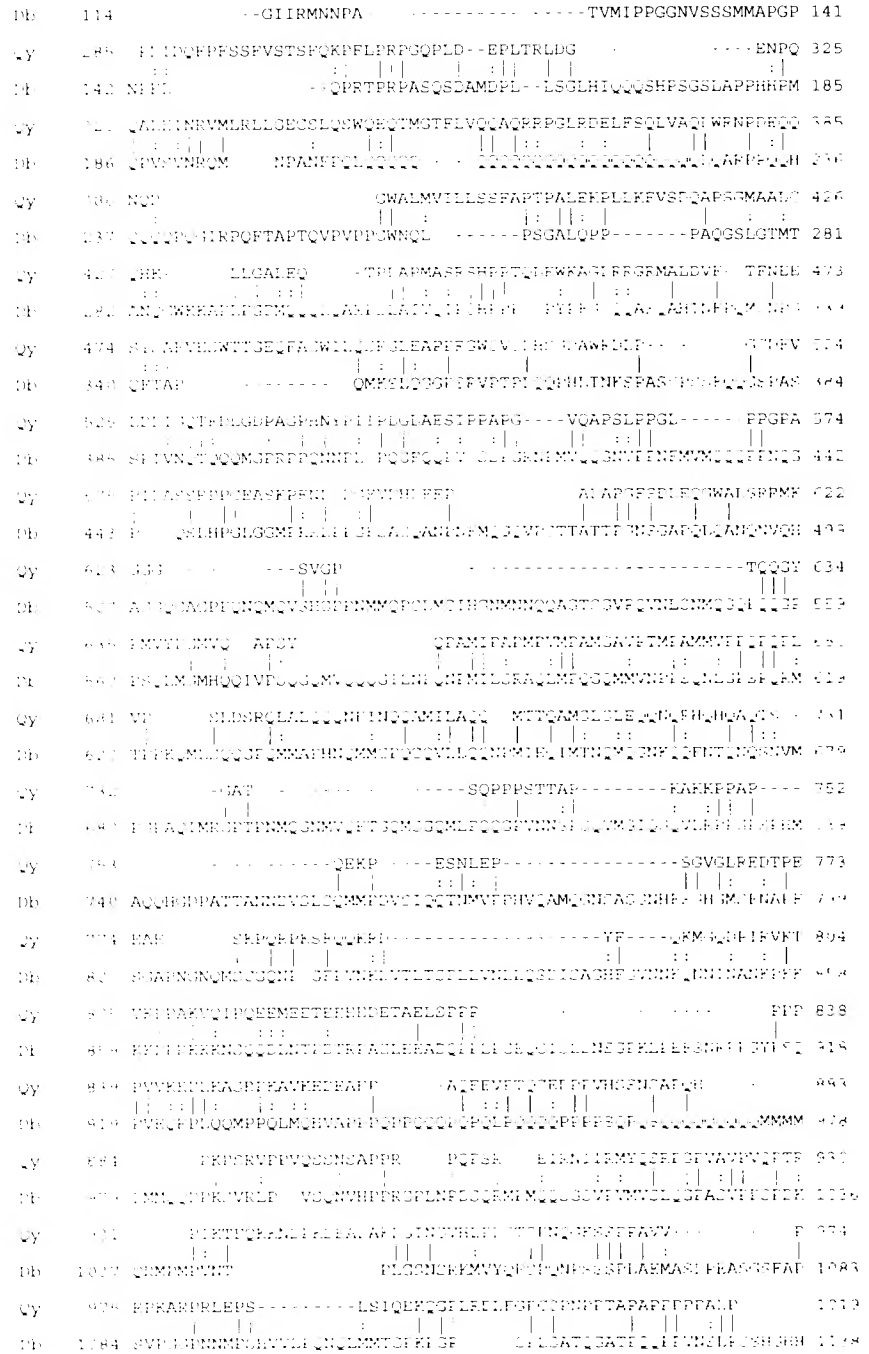
Query Match 3.8%; Score 267.5; DB 10; Length 1404;
Best Local Similarity 21.6%; Pred. No. 1.2e-06;
Matches 183; Conservative 79; Mismatches 256; Indels 221; Gaps 37;

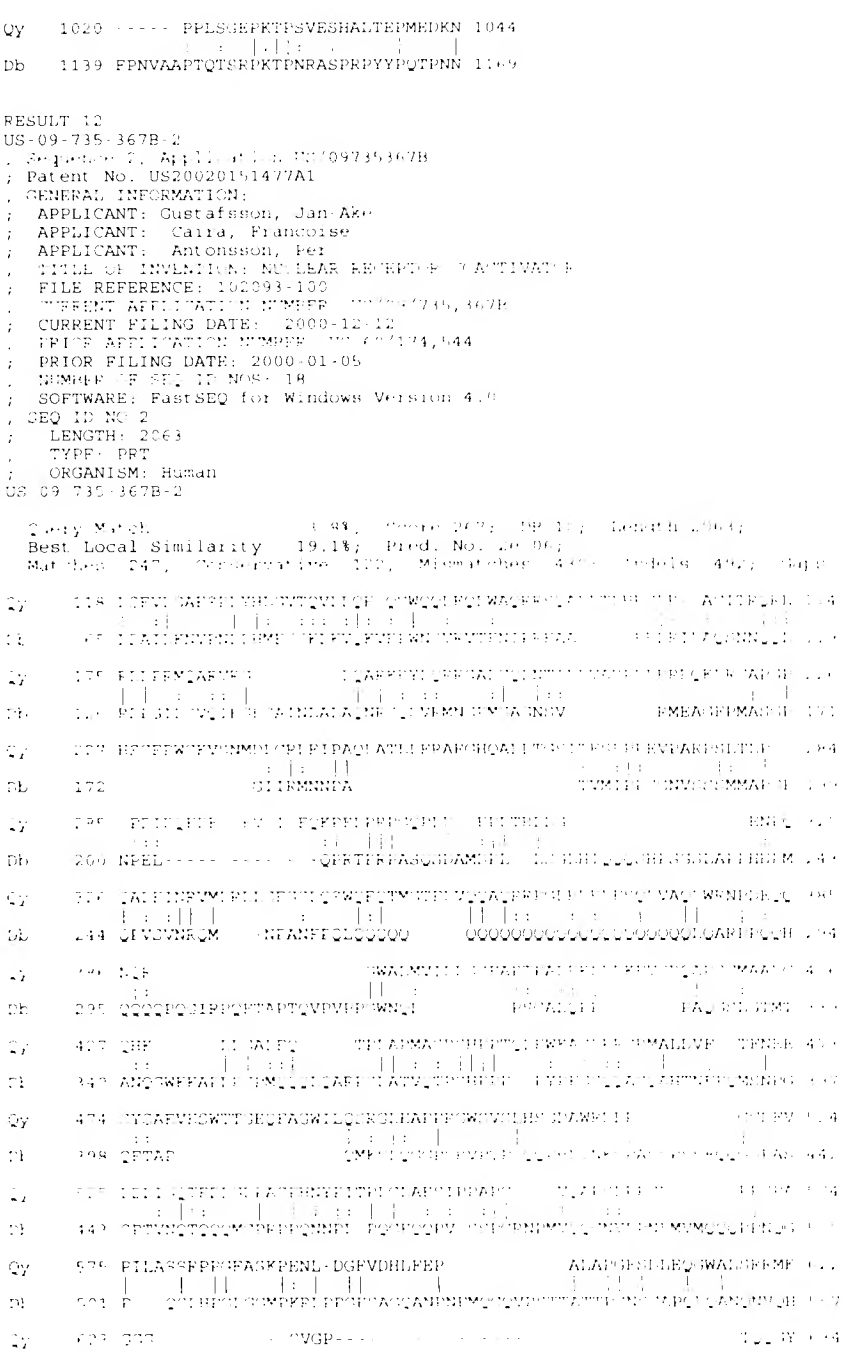


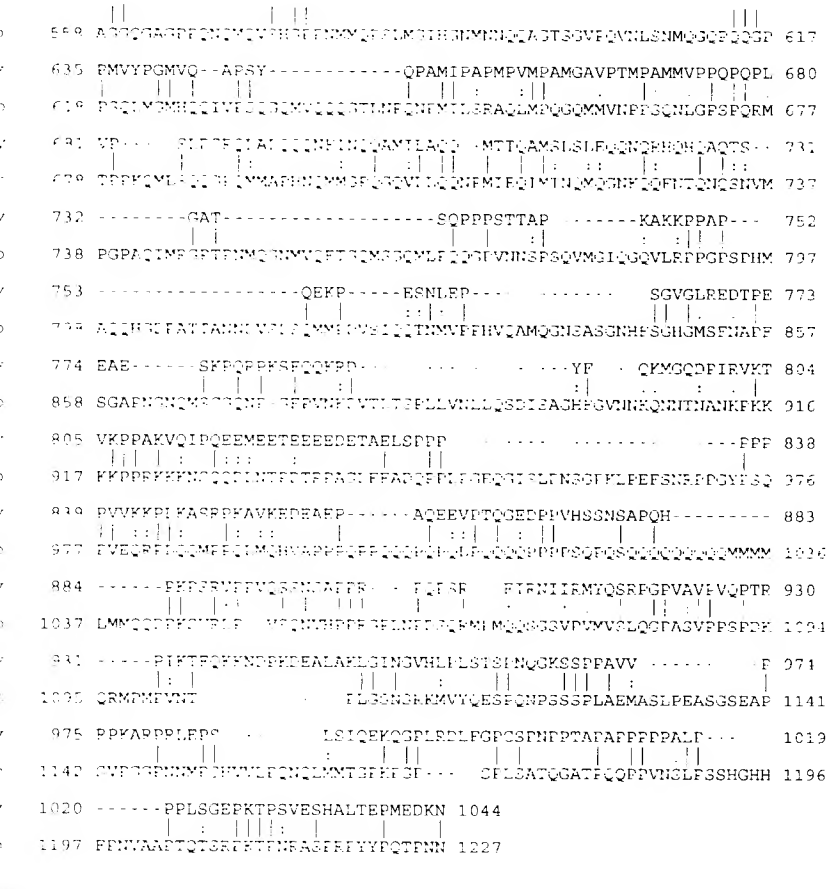
RESULT 11
US-09-735-367B-3
; Sequence 3, Application US/09735367B
; Patent No. US20020151477A1
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Jan-Ake
; APPLICANT: Caira, Francoise
; APPLICANT: Antonsson, Per
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR
; FILE REFERENCE: 102093-100
; CURRENT APPLICATION NUMBER: US/09/735,367B
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/174,544
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2005
; TYPE: PRT
; ORGANISM: mammal
US-09-735-367B-3

Query Match 3.8%; Score 267; DB 10; Length 2005;
Best Local Similarity 19.1%; Pred. No. 1.9e-06;
Matches 247; Conservative 122; Mismatches 430; Indels 492; Gaps 58;



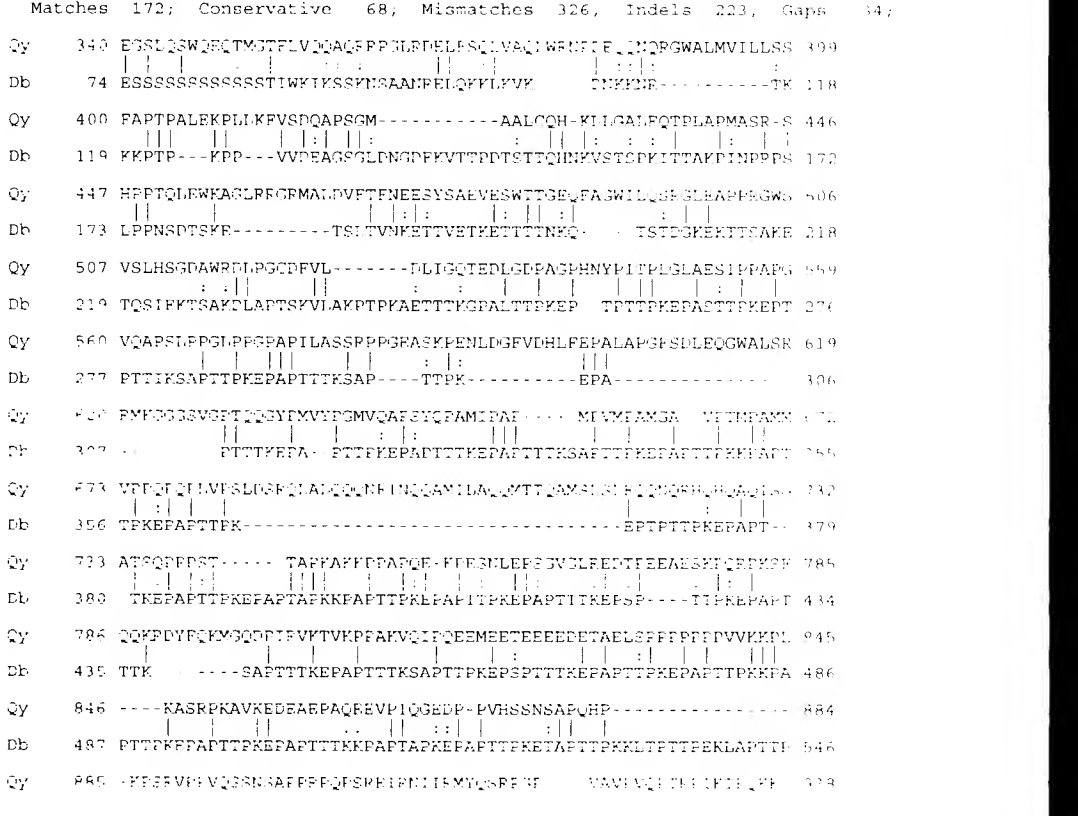


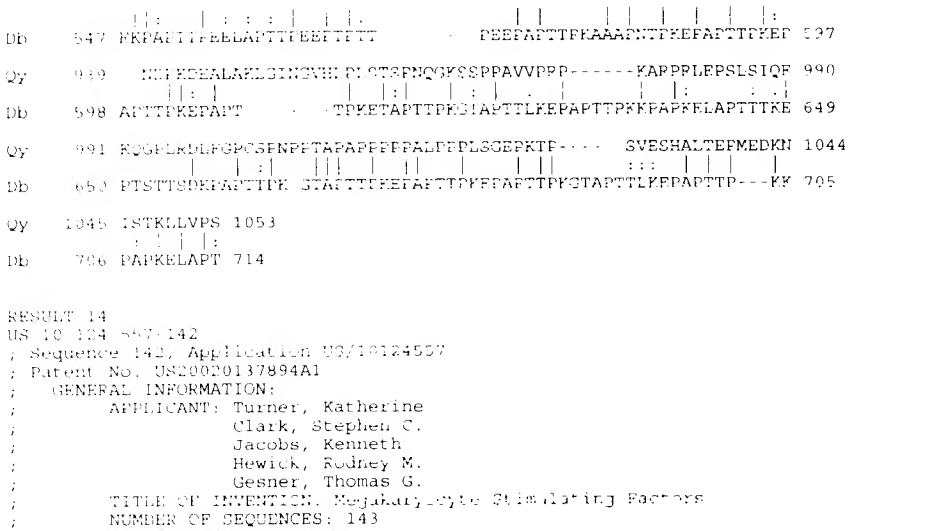




ESULT 13
-10-124-557-58
Sequence 58, Application US/10124557
Patent No. US20020137894A1
GENERAL INFORMATION
APPLICANT: Turner, Katherine
Clark, Stephen C
Jacobs, Kenneth
Hewick, Rodney M
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS
ADDRESSEE: Genetics Institute, Inc
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 1049 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-124-557-58

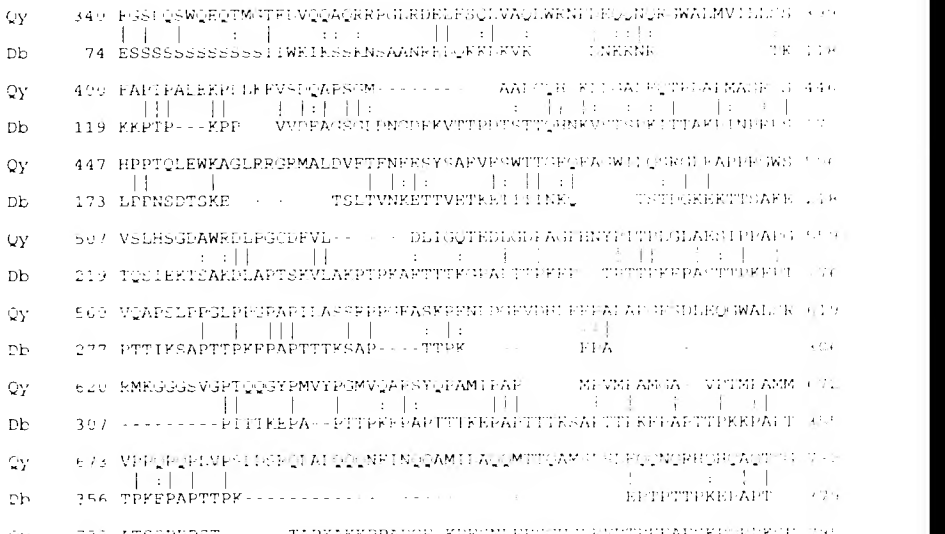
Query Match 3.8%; Score 264.5; DB 12; Length 1049;
Best Local Similarity 21.8%; Pred. No. 1.2e-06;
Matches 172; Conservative 68; Mismatches 326; Indels 223; Gaps 34;





RESULT 14
US 10/124,557-142
; Sequence 142, Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/540,114
; FILING DATE: 19-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS
; LENGTH: 1313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 142.
US 10/124,557-142

Query Match 3.8%; Score 264.5; DB 12; Length 1313;
Best Local Similarity 21.8%; Pred. No. 1.6e-14;
Matches 172; Conservative 68; Mismatches 320; Indels 223; Gaps 34;



RESULT 15
US-10-124,557-50
; Sequence 50, Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge

STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 28-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION.
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO. 50.
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION. SEQ ID NO: 50:
S-10-124-557-50

Query Match 3.8%, Score 264.5, DB 12, Length 1314,
Best Local Similarity 21.8%, Pred. No. 1.6e-06,
Matches 170; Conservative 68; Mismatches 320; Indels 223; Gaps 34;

```
Qy    340 EGSLQSWQEQTMGTFLVQQAQRRPGLRDELFSQLVAQLWRNPDEQQNQRGWALMVILLSS 399
Db     76 EGSSQSSSQSSPTIWKIKSSKNSAANRELQKKLKVK- ---DNKKIR----------TK 119

Qy    400 EAPTPAIEKKLLKFVSDQAPSGM----------AALCQH-KLLGALEQTPLAPMASR-S 446
Db    120 KKPTP---KPP---VVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPPPS 173

Qy    447 HDPTQLEWKAGLKRQRMALDVFTFNEECYGAEVESWTTGEQFAGWILQSRGLEAPPRGWS 506
Db    174 LPPNSDTSKE ------TSLTVNKETTVETKETTTTNKQ TSTDGKEKTTSAKE 219

Qy    507 VSIHSGDAWFDLPGCDFVL-------DLIGQTEDLGDPAGPHNYPITPLGLAESIPPAPG 559
Db    220 EQSIFKTSAKDLAPTSKVLAKPTPKAETTTVGPALTTPKEP- TPTTPKEPASTTPKEPT 277

Qy    560 VQAPSLPPGLPPGPAPILASSKPPGEASKPENLDGFVDHLFEPALAPGFSDLEQGWALSR 619
Db    278 PTTIKSAPTTPKEPAPTTTKSAP----TTPK----------EPA--------------- 307

Qy    620 EMKGGGSVGFTQQGYEMVYPGMVQAFSYQPAMIPAP-----MPVMPAMGA--VPTMPAMM 672
Db    308 ---------PTTTKEPA--PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPT 356

Qy    673 VEEIPGEIVEIIKPQLALGQNPINQQAMILAQQMTTQAMSLSLPQQNQPHQHQAQTSG 732
Db    357 TPKEPAPTTPK---------------------------------EPTPTTPKEPAPT-- 380

Qy    733 ATSQPPPST-- TAPKAKPPPAPQF-KPPSNLEPSGVGLPEDTPEEAESKPQPPKSF 785
Db    381 -TKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSP----TTPKEPAPT 435
```

```
Qy    786 QQKPDVPQKMGQDPIPVKTVKPPAKVQIPQEEMEETEEFEDFTAELSPPPPPPPVVKKPL 845
Db    436 TTK---------SAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPA 487

Qy    846 ----KASRPKAVKEDEAEPAQEEVPTQGEDP-PVHSSNSAPQHP---------------- 884
Db    488 PTTPKEPAPTTPKEPAPTITKPAPTAPKEPAPTTPKETAPTTPKKLIPTTPEPLAPTTP 547

Qy    885 -KPSRVPPVQSSNSAPPRPQPSPEIPNIIPMYQSPPGF VAVPVQPTPPIPTPQPP 938
Db    548 EKPAPTTPEELAPTTPEEPTPTT- - - PEEPAPTTPKAAAAPNTPKEPAPTTPKEP 598

Qy    939 --NDPKDEALAKLGIPGVHLPLSTSPNQGPQSPPAVVPPI-- PAPIPLEPILTIQI 990
Db    599 APTTPKEPAPT- - - -TPKETAPTTPKGTAPTTLKEPAPTTPKPPAPKELAPTTTPK 650

Qy    991 KQGPLPDLPGPCSPNPPTAPAPPPPPALPPPLSGEPKTP------SVESHALTPPMPDKN 1044
Db    651 PTSITSDKPAPTTPK-GIAPTTPKEPAPTTPKEPAPTTPKGTAPTTDKEPAPTTP--PK 706

Qy   1045 ISTKLLVPS 1053
Db    707 PAPKELAPT 715
```

Search completed: January 10, 2003, 02:29:03
Job time : 113.409 secs

GenCore version 5.1.3

M nucleic - nucleic search, using sw model

un on: January [illegible] 2003, [illegible] ; Search time 15302 Seconds
(without alignments)
119[illegible] Million cell updates/sec

itle: US-09-803-126-2
erfect score: 6293
equence: 1 [illegible] .aaaaaaaaaaaaaaaaaaaa 6293

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

earched: 2054640 seqs, 14551402878 residues

otal number of hits satisfying chosen parameters: 4109280

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 1%
Maximum Match 100%
Listing first 45 summaries

atabase: GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: [illegible]
21: em_or:*
22: [illegible]
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

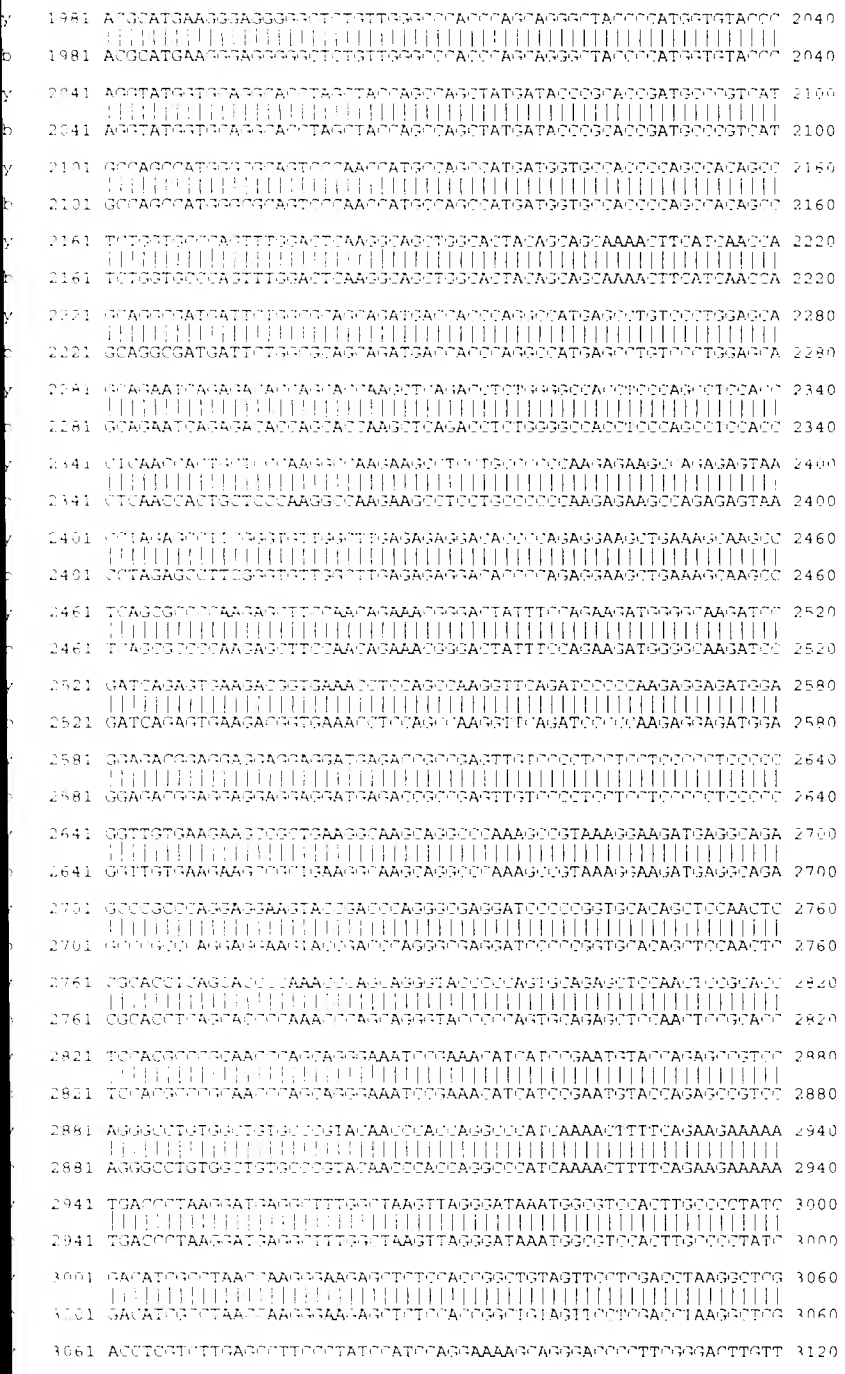
SUMMARIES

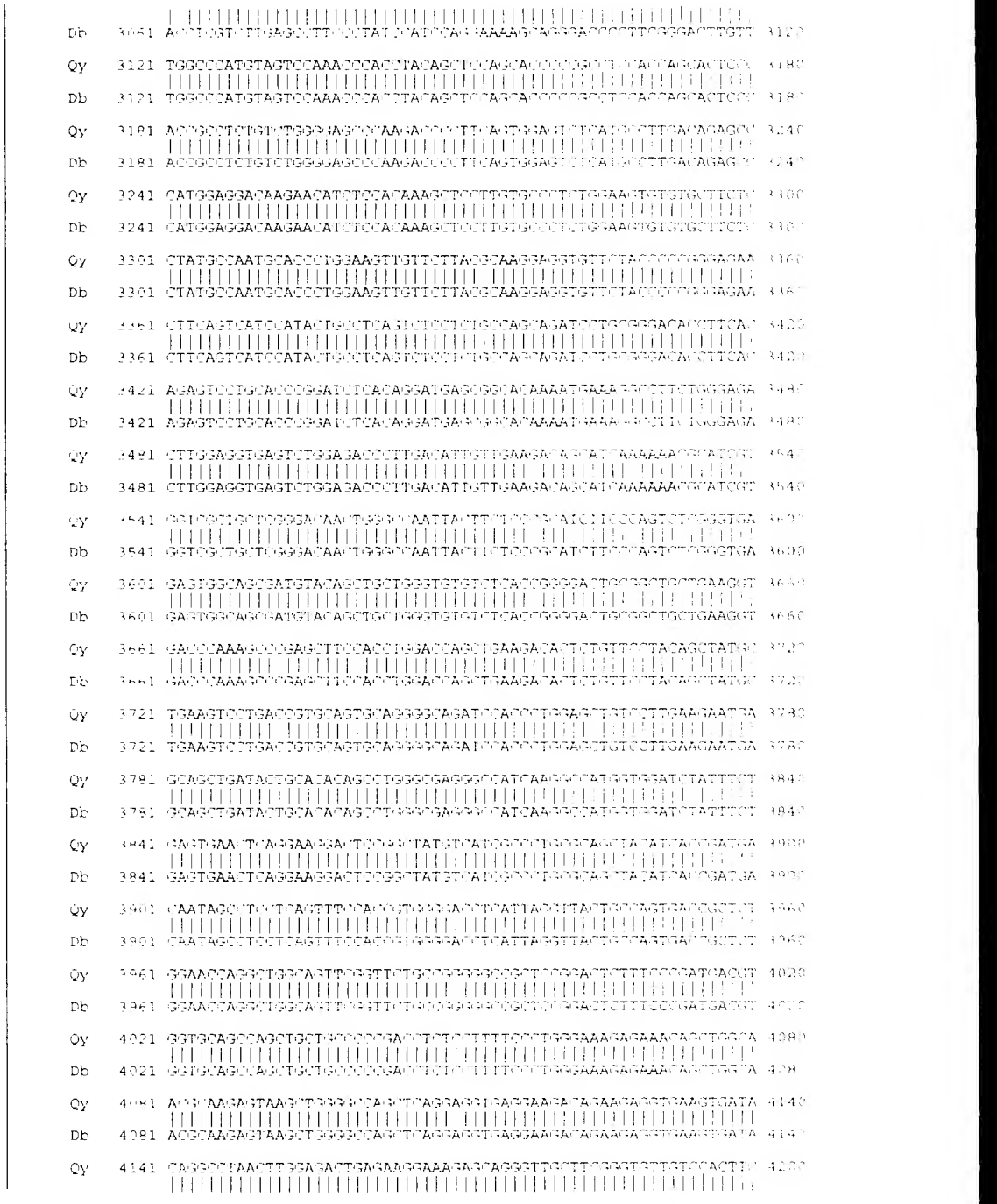
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	6293	100.0	6293	6	AX253079	AX253079 Sequence
	2	4351	69.2	4375	6	AX25[illegible]	AX[illegible] Sequence
	3	2384.2	37.9	6004	9	AB058686	AB058686 Homo sapi
	4	2246	35.7	4602	9	AF418292	AF418292 Homo sapi
	5	2055.4	32.7	7925	9	AF418290	AF418290 Homo sapi
	6	1954.4	31.1	7515	9	AF418286	AF418286 Homo sapi
	7	1949.6	31.0	4208	9	AF418293	AF418293 Homo sapi
	8	1787.6	28.4	4174	6	AX253082	AX253082 Sequence
	9	1495.6	23.9	1547	10	BC024506	BC024506 Mus muscu
	10	1490	23.7	3780	6	AX253084	AX253084 Sequence
	11	1119.4	17.8	5709	9	AF418289	AF418289 Homo sapi
	12	1110.6	17.6	2372	9	AK097896	AK097896 Homo sapi
	13	936.8	14.9	2347	9	BC027875	BC027875 Homo sapi
	14	872.6	13.9	1891	9	AK026339	AK026339 Homo sapi
	15	861	13.7	5424	9	AF418288	AF418288 Homo sapi
	16	822.6	13.1	1767	9	HSM803107	AL713730 Homo sapi
	17	721	11.5	147804	10	AL645647	AL645647 Mouse DNA
c	18	721	11.5	260310	2	AL645852	AL645852 Mus muscu
	19	612.2	9.7	1998	9	AK096869	AK096869 Homo sapi
	20	577	9.2	4802	9	AF418291	AF418291 Homo sapi
	21	544.2	8.6	2099	9	AK025259	AK025259 Homo sapi
c	22	474.6	7.5	180356	2	AC123410	AC123410 Rattus no
	23	474.6	7.5	200832	2	AC130970	AC130970 Rattus no
	24	371.6	5.9	4462	9	AF418287	AF418287 Homo sapi
	25	311.2	4.9	11876	9	AF144094	AF144094 Homo sapi
	26	291.4	4.6	11769	10	AF144095	AF144095 Mus muscu
	27	286.6	4.6	5351	10	AF053130	AF053130 Mus muscu
	28	263.2	4.2	180956	2	AC123410	AC123410 Rattus no
c	29	262.2	4.2	200832	2	AC130970	AC130970 Rattus no
	30	195.6	3.1	170056	2	AC087643	AC087643 Homo sapi
	31	195.6	3.1	178515	2	AC015551	AC015551 Homo sapi
	32	184.2	2.9	217584	2	AC087749	AC087749 Homo sapi
c	33	166.2	2.6	215434	2	AC019214	AC019214 Homo sapi
c	34	111.6	1.8	170056	2	AC087643	AC087643 Homo sapi
c	35	102	1.6	178515	2	AC015551	AC015551 Homo sapi
	36	86.2	1.4	191184	2	AC120677	AC120677 Rattus no
	37	86.2	1.4	195469	2	AC129460	AC129460 Rattus no
	38	84.6	1.3	88406	10	AF144093	AF144093 Mus muscu
	39	84.6	1.3	261189	2	AL596286	AL596286 Mus muscu
	40	81.6	1.3	125020	9	AF429215	AF429215 Homo sapi
	41	76.6	1.2	95255	9	AF051976	AF051976 Homo sapi
c	42	76.6	1.2	144448	2	AC109515	AC109515 Homo sapi
c	43	76.6	1.2	164429	2	AC087164	AC087164 Homo sapi
c	44	76.6	1.2	173749	2	[illegible]	[illegible] Homo sapi
	45	76.4	1.2	3110	9	HSM803101	AL713794 Homo sapi

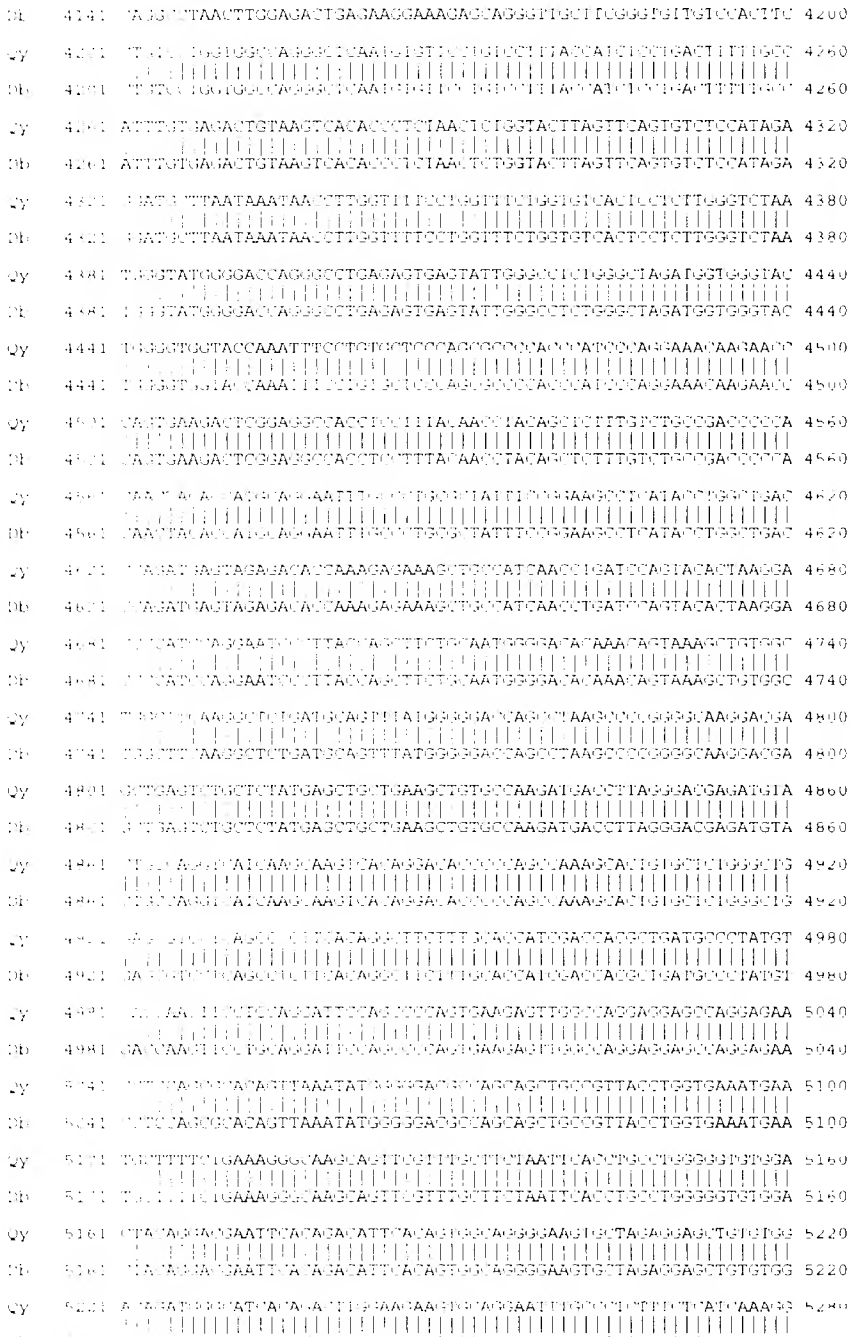
ALIGNMENTS

RESULT 1
AX253079
LOCUS AX253079 6293 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 2 from Patent WO0168866.
ACCESSION AX253079
VERSION AX253079.1 GI:15986277
KEYWORDS .
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 6293)
AUTHORS Brooks,A.R., Deng,G.G. and Rubanyi,G.M.
TITLE Estrogen regulated unconventional myosin-related protein:
compositions and methods of use

```
JOURNAL    Patent: WO 0168866-A 2 20-SEP-2001;
           SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1..6293
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
                     /note="Mouse myosin related protein (MRP) variant 1 cDNA"
BASE COUNT     1433 a   1917 c   1733 g   1210 t
ORIGIN

  Query Match          100.0%;  Score 6293;  DB 6;  Length 6293;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 6293;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGTTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60

Qy         61 CTCAGCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCAGCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120

Qy        121 GCTCATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGGCTGGGACTGAGCAAAACAAACC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTCATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGGCTGGGACTGAGCAAAACAAACC 180

Qy        181 CACATTCGCCTCTCGATTCCAGCAGACCCTGGGTGACTTGCTAGCTCGGCTAGGCAGCAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACATTCGCCTCTCGATTCCAGCAGACCCTGGGTGACTTGCTAGCTCGGCTAGGCAGCAG 240

Qy        241 GGACCATGTCTACGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACCATGTCTACGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300

Qy        301 GGACCTGGGCATGTGGCAGAGCAGCTGCGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGACCTGGGCATGTGGCAGAGCAGCTGCGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360

Qy        361 CCGGAGTACCCACTTCCCCGTGCGAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGAGTACCCACTTCCCCGTGCGAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420

Qy        421 CCTGGGGTCAGGGAGACAGAAAGCTGCCTCTGACCAGGAGAGGTGTGGTGCCATCCTCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGGGGTCAGGGAGACAGAAAGCTGCCTCTGACCAGGAGAGGTGTGGTGCCATCCTCAG 480

Qy        481 TGAAGTGCTGGGGGCAGAGCTACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGAAGTGCTGGGGGCAGAGCTACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCA 540

Qy        541 GGAACAGGGCTGGCAGCAGCTAGAACAGCTGTGGGCTCAGCGGCGCTCACAGGCCCTGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAACAGGGCTGGCAGCAGCTAGAACAGCTGTGGGCTCAGCGGCGCTCACAGGCCCTGCT 600

Qy        601 CACTCTGCACCGTGGCCTTGGAGCCTGTATCACCGGCAGCGCCTCCGTCTCCTGCCCCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACTCTGCACCGTGGCCTTGGAGCCTGTATCACCGGCAGCGCCTCCGTCTCCTGCCCCG 660

Qy        661 GATGCAGGCTCGTGTGCGTGGGCTCCAGGCCAGGAAGGCATATCTCCAGCGGAGGTCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGCAGGCTCGTGTGCGTGGGCTCCAGGCCAGGAAGGCATATCTCCAGCGGAGGTCAGC 720

Qy        721 TCTGGGACAGCTGAACACCATTCTCCTAGTGGCCCGGCCCCTGCTCCGGAGACGACAGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTGGGACAGCTGAACACCATTCTCCTAGTGGCCCGGCCCCTGCTCCGGAGACGACAGAA 780

Qy        781 GCTACGGTGTGCCCCTGGTCCGCACAGCGGGGAGCCCTGGGGGAAAGTGTCAAATATGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCTACGGTGTGCCCCTGGTCCGCACAGCGGGGAGCCCTGGGGGAAAGTGTCAAATATGGA 840

Qy        841 CCTGGGTCGCTTAGAGATCCCCGCCCAGCTGGCTACTCTGCTGGAGAGGCGGAAGGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGGTCGCTTAGAGATCCCCGCCCAGCTGGCTACTCTGCTGGAGAGGCGGAAGGCCA 900

Qy        901 CCAGGCTTTGCTGACGGGGAGCATCACAGAGTCCCTG[illegible] 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCAGGCTTTGCTGACGGGGAGCATCACAGAGTCCCTG[illegible] 960

Qy        961 CAGCCTGACTCTCCCTCCAGACATTGACCAGTTTCCCTTCTCCAGTTTTGTATCCACCA[illegible] 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCCTGACTCTCCCTCCAGACATTGACCAGTTTCCCTTCTCCAGTTTTGTATCCACCA[illegible] 1020

Qy       1021 CTTTCAGAAGCCATTTCTGCCTCGACCAGGGCAGCCAC[illegible] 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTTCAGAAGCCATTTCTGCCTCGACCAGGGCAGCCAC[illegible] 1080

Qy       1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAACAGG[illegible] 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAACAGG[illegible] 1140

Qy       1141 GGAAGGATCTCTGCAGTCCTGGCAAGAGCAGACCAT[illegible] 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAAGGATCTCTGCAGTCCTGGCAAGAGCAGACCAT[illegible] 1200

Qy       1201 CCAGCGACGGCCGGGACTCCGAGATGAGCTCTTAG[illegible] 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCAGCGACGGCCGGGACTCCGAGATGAGCTCTTAG[illegible] 1260

Qy       1261 CAACCCAGATGAGCAACAGAATCAGGTGGC[illegible] 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CAACCCAGATGAGCAACAGAATCAGGTGGC[illegible] 1320

Qy       1321 CTTTGCTCCCACACCTGCCCTGGAGAAGCCACTG[illegible] 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTTTGCTCCCACACCTGCCCTGGAGAAGCCACTG[illegible] 1380

Qy       1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAGCTGTTAG[illegible] 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAGCTGTTAG[illegible] 1440

Qy       1441 GGCTCCCATGGCTTCGAGGTCCCACCCACCCACACAA[illegible] 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GGCTCCCATGGCTTCGAGGTCCCACCCACCCACACAA[illegible] 1500

Qy       1501 TCGGGGCCGCATGGCGCTGGATGTGTTCACATTCAACGA[illegible] 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCGGGGCCGCATGGCGCTGGATGTGTTCACATTCAACGA[illegible] 1560

Qy       1561 GGAATCCTGGACCACGGGAGAGCAGTTTGCAGGG[illegible] 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAATCCTGGACCACGGGAGAGCAGTTTGCAGGG[illegible] 1620

Qy       1621 GGCGCCCCCTCGTGGCTGGTCTGTGTCACTGCATTC[illegible] 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGCGCCCCCTCGTGGCTGGTCTGTGTCACTGCATTC[illegible] 1680

Qy       1681 TGGCTGTGACTTTGTGTTGGACCTAATAGG[illegible] 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGGCTGTGACTTTGTGTTGGACCTAATAGG[illegible] 1740

Qy       1741 TCCCCACAACTACCCCATCACTCCTCTTGGTTTA[illegible] 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCCCCACAACTACCCCATCACTCCTCTTGGTTTA[illegible] 1800

Qy       1801 TGTCCAGGCTCCTTCCCTGCCCCCAGGAC[illegible] 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGTCCAGGCTCCTTCCCTGCCCCCAGGAC[illegible] 1860

Qy       1861 CAGCCGCCCTCCGGGCGAGGCCAGTAAGCCTGAGAA[illegible] 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CAGCCGCCCTCCGGGCGAGGCCAGTAAGCCTGAGAA[illegible] 1920

Qy       1921 CTTTGAACCAGCGCTCGCTCCGGGTTTCAGTGATCTGGAA[illegible] 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTTTGAACCAGCGCTCGCTCCGGGTTTCAGTGATCTGGAA[illegible] 1980
```

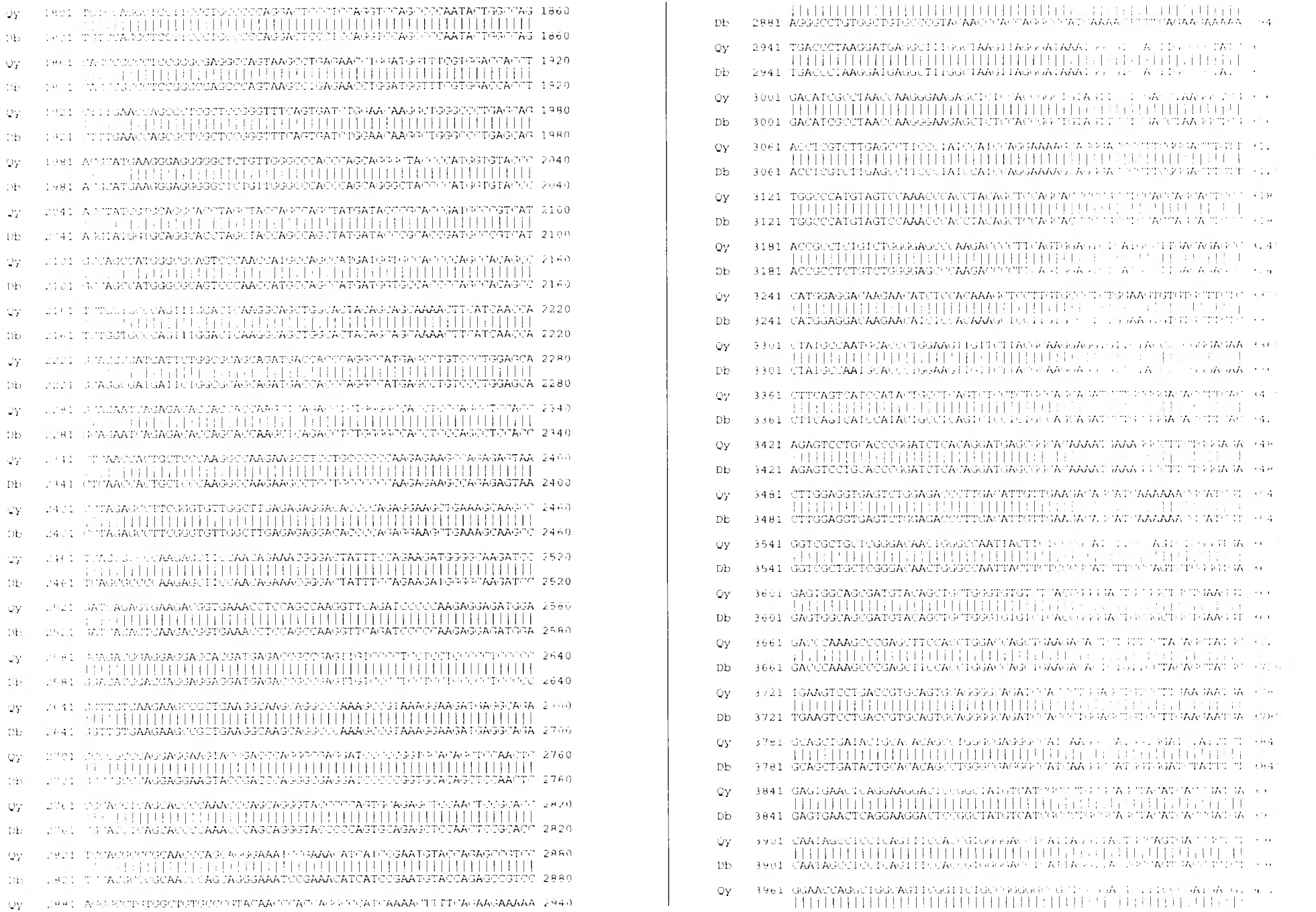




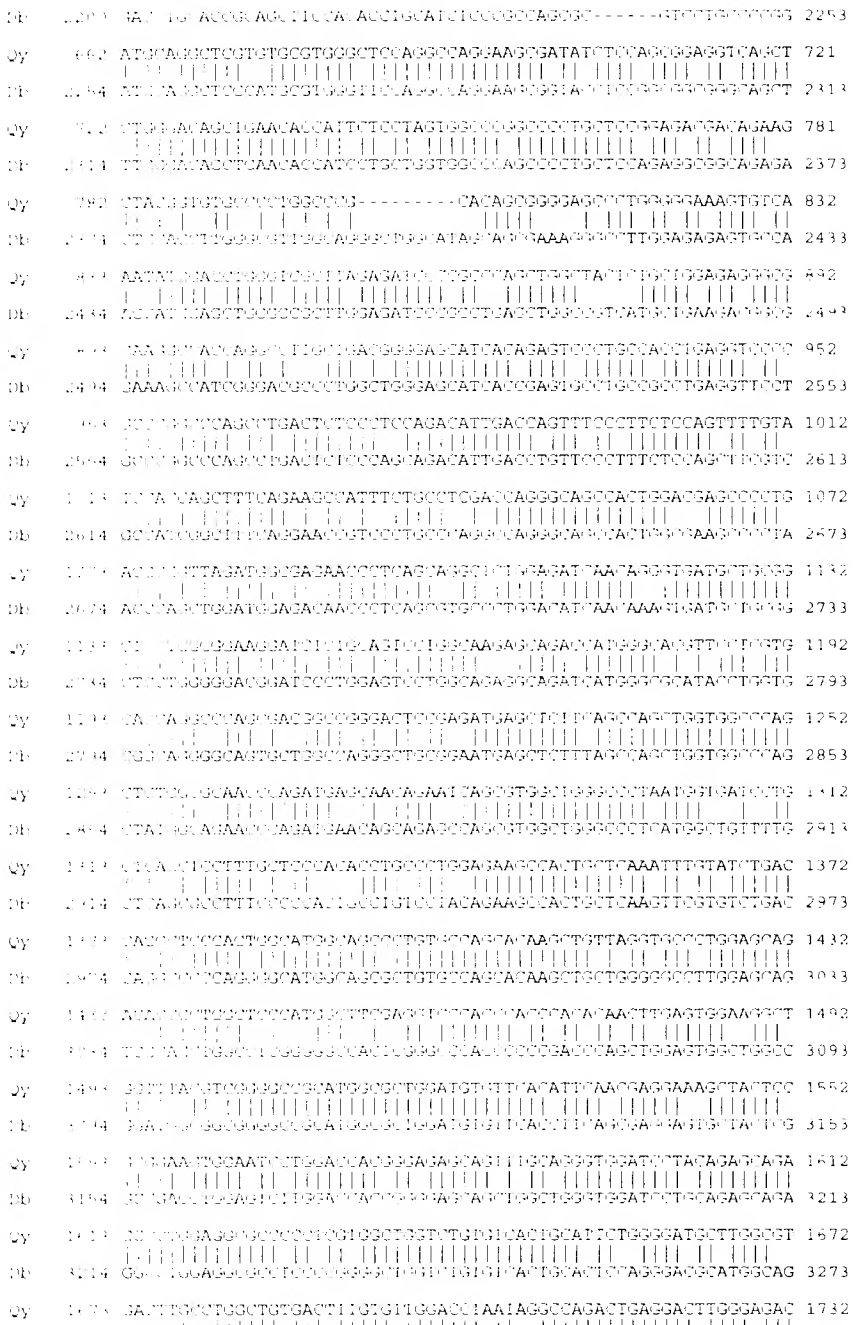


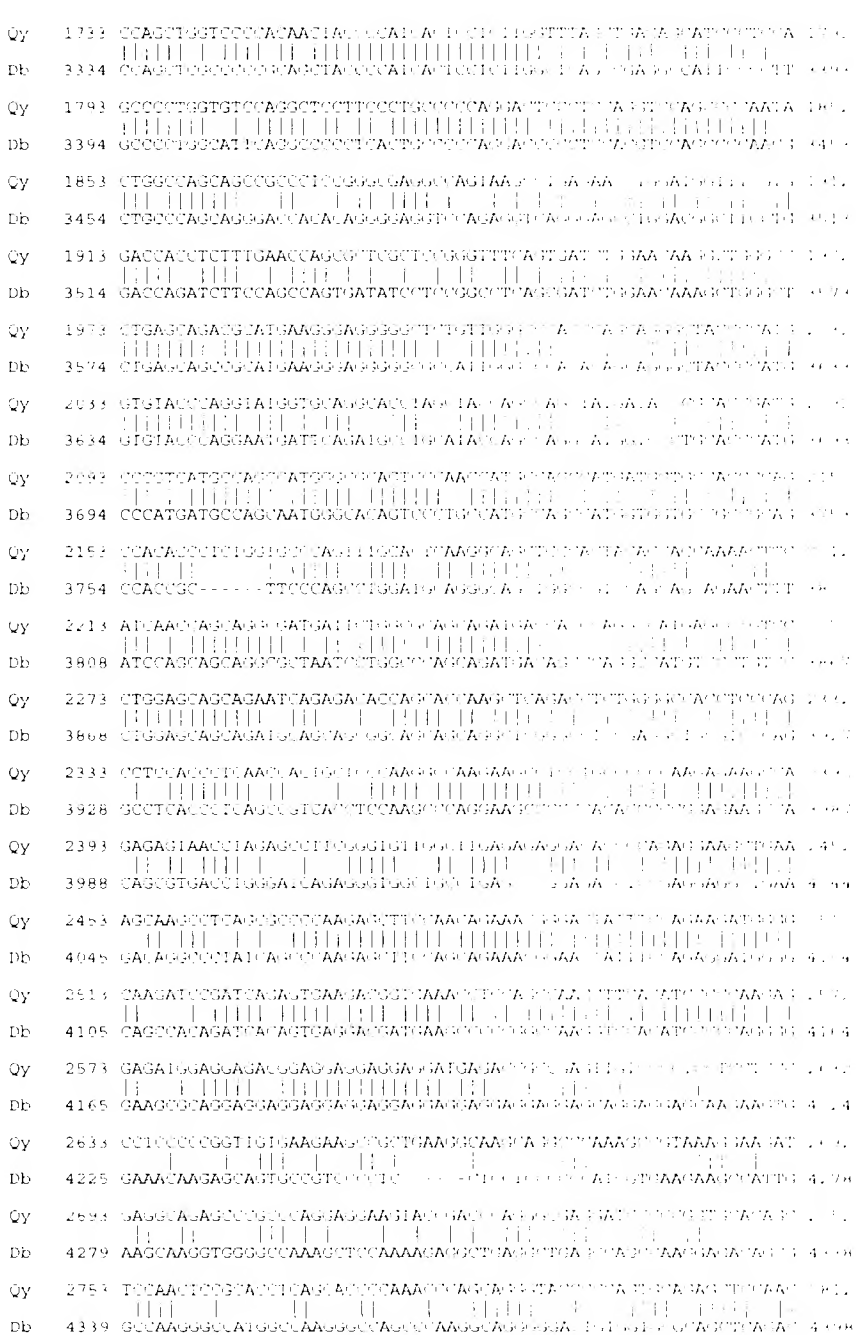
RESULT 2
AX253080
LOCUS AX253080 4479 bp DNA

```
EFINITION   Sequence 3 from Patent WO0168866.
CCESSION    AX253080
ERSION      AX253080.1  GI:[illegible]
EYWORDS     .
OURCE       Mus sp.
 ORGANISM   Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
EFERENCE    1  (bases 1 to 4375)
 AUTHORS    Brooks,A.R., Deng,G.G. and Rubanyi,G.M.
 TITLE      Estrogen-regulated unconventional myosin related protein:
            compositions and methods of use
 JOURNAL    Patent: W [illegible] SEP 2001,
            SCHERING AKTIENGESELLSCHAFT (DE)
EATURES              Location/Qualifiers
    source           1..4375
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
                     /note="Mouse myosin related protein (MRP) variant 2 cDNA"
ASE COUNT      966 a   1368 c   1216 g    825 t
RIGIN

 Query Match            [illegible]%;  Score 4352;  DB 6;  Length 4375;
 Best Local Similarity  100.0%;  Pred. No. 0;
 Matches 4352;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      1 CGCTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1 CGCTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60

     61 [illegible]GGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     61 [illegible]GGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120

    121 GCT[illegible]AGCCTGTTCCAAGAAG[illegible]AGGCTGGGACTGAGCAAAACAAACC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    121 GCT[illegible]AGCCTGTTCCAAGAAG[illegible]AGGCTGGGACTGAGCAAAACAAACC 180

    181 CACATT[illegible]GATT[illegible]AGA[illegible]CTGGGTGACTTGCTAGCTCGGCTAGGCAGCAG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    181 CACATT[illegible]GATT[illegible]AGA[illegible]CTGGGTGACTTGCTAGCTCGGCTAGGCAGCAG 240

    241 [illegible]CATG[illegible]TAGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    241 [illegible]CATG[illegible]TAGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300

    301 [illegible]GCAGCTGGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    301 [illegible]GCAGCTGGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360

    361 CCGGAGTACCCA[illegible]GAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    361 CCGGAGTACCCA[illegible]GAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420

    421 [illegible]AGAGAGAAAGCTGCCCTCTGACCAGGAGAGGTGTGGTGCCATCCTCAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    421 [illegible]AGAGAGAAAGCTGCCCTCTGACCAGGAGAGGTGTGGTGCCATCCTCAG 480

    481 TGAAG[illegible]AGAGTCACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCA 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    481 TGAAG[illegible]AGAGTCACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCA 540

    541 [illegible]GAAACAGCTTGTGGGCTCAGCGGCGCTCACAGGCCCTGCT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    541 [illegible]GAAACAGCTTGTGGGCTCAGCGGCGCTCACAGGCCCTGCT 600

    601 [illegible]GGAGCCTGTATCACCCGGCAGCGCCTCCGTCTCCTGCCCCG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    601 [illegible]GGAGCCTGTATCACCCGGCAGCGCCTCCGTCTCCTGCCCCG 660

    661 [illegible]CCAGGCCAGGAAGCGATATCTCCAGCGGAGGTCAGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    661 [illegible]CCAGGCCAGGAAGCGATATCTCCAGCGGAGGTCAGC 720

Qy  721 TCTGGGACAGCTGAACACCATTCTCCTAGTGG[illegible]GGAGACGACAGAA 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TCTGGGACAGCTGAACACCATTCTCCTAGTGG[illegible]GGAGACGACAGAA 780

Qy  781 GCTACGGTGTGCCCCTGGCCCGCACAGCGGGGAG[illegible]GGGAAAGTGTCAAATATGGA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GCTACGGTGTGCCCCTGGCCCGCACAGCGGGGAG[illegible]GGGAAAGTGTCAAATATGGA 840

Qy  841 CCTGGGTCGCTTAGAGATCCCCGCCCAGCTGGCTACTCTGCTGGAGAGGGCGGAAG[illegible]CA 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CCTGGGTCGCTTAGAGATCCCCGCCCAGCTGGCTACTCTGCTGGAGAGGGCGGAAG[illegible]CA 900

Qy  901 CCAGGCCTTGCTGACGGGGAGCATCACAGAGTCCCTGCCACCTGAGGT[illegible]GGCC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CCAGGCCTTGCTGACGGGGAGCATCACAGAGTCCCTGCCACCTGAGGT[illegible]GGCC 960

Qy  961 CAGCCTGACTCTCCCTCCAGACATTGACCAGTTTCCCTTCTCCAGTTTTGTATCCACCAG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CAGCCTGACTCTCCCTCCAGACATTGACCAGTTTCCCTTCTCCAGTTTTGTATCCACCAG 1020

Qy 1021 CTTTCAGAAGCCATTTCTGCCTCGACCAGGGCAGCCACTGGACGAGCCCCTGACGCGGTT 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CTTTCAGAAGCCATTTCTGCCTCGACCAGGGCAGCCACTGGACGAGCCCCTGACGCGGTT 1080

Qy 1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAACAGGGTGATGCTGCGGCTCCTGGG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAACAGGGTGATGCTGCGGCTCCTGGG 1140

Qy 1141 GGAAGGATCTCTGCAGTCCTGGCAAGAGCAGACCATGGGCA[illegible] 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GGAAGGATCTCTGCAGTCCTGGCAAGAGCAGACCATGGGCA[illegible] 1200

Qy 1201 CCAGCGACGGCCGGGACTCCGAGATGAGCTCTTCAGCCAG[illegible] 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CCAGCGACGGCCGGGACTCCGAGATGAGCTCTTCAGCCAG[illegible] 1260

Qy 1261 CAACCCAGATGAGCAACAGAATCAGGGTGGCTGGGCCCTAATGGTGAT[illegible] 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CAACCCAGATGAGCAACAGAATCAGGGTGGCTGGGCCCTAATGGTGAT[illegible] 1320

Qy 1321 CTTTGCTCCCACACCTGCCCTGGAGAAGCCACTGCTCAAATTTGTATCTGACCAGGCTCC 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CTTTGCTCCCACACCTGCCCTGGAGAAGCCACTGCTCAAATTTGTATCTGACCAGGCTCC 1380

Qy 1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAGCTGTTAGGTG[illegible]CT 1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAGCTGTTAGGTG[illegible]CT 1440

Qy 1441 GGCTCCCATGGCTTCGAGGTCCCACCCACCCACACAACTTGAGTGGAAGG[illegible]TGGTTTACG 1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GGCTCCCATGGCTTCGAGGTCCCACCCACCCACACAACTTGAGTGGAAGG[illegible]TGGTTTACG 1500

Qy 1501 TCGGGGCCGCATGGCGCTGGATGTGTTCACATTCAACGAGGAAAGCTACTCCGGGGAACT 1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 TCGGGGCCGCATGGCGCTGGATGTGTTCACATTCAACGAGGAAAGCTACTCCGGGGAACT 1560

Qy 1561 GGAATCCTGGACCACGGGAGAGCAGTTTGCAGGGTGGATCCTACAGAGCAGAG[illegible]TGGA 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 GGAATCCTGGACCACGGGAGAGCAGTTTGCAGGGTGGATCCTACAGAGCAGAG[illegible]TGGA 1620

Qy 1621 GGCGCCCCCTCGTGGCTGGTCTGTGTCACTGCATTCTGGGGATG[illegible] 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 GGCGCCCCCTCGTGGCTGGTCTGTGTCACTGCATTCTGGGGATG[illegible] 1680

Qy 1681 TGGCTGTGACTTTGTGTTGGACCTAATAGGCCAGACTGAGGACTTTGGAGA[illegible] 1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TGGCTGTGACTTTGTGTTGGACCTAATAGGCCAGACTGAGGACTTTGGAGA[illegible] 1740

Qy 1741 TCCCCACAACTACCCCATCACTCCTCTTGGTTTAGCTGAGA[illegible]TGG 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 TCCCCACAACTACCCCATCACTCCTCTTGGTTTAGCTGAGA[illegible]TGG 1800
```



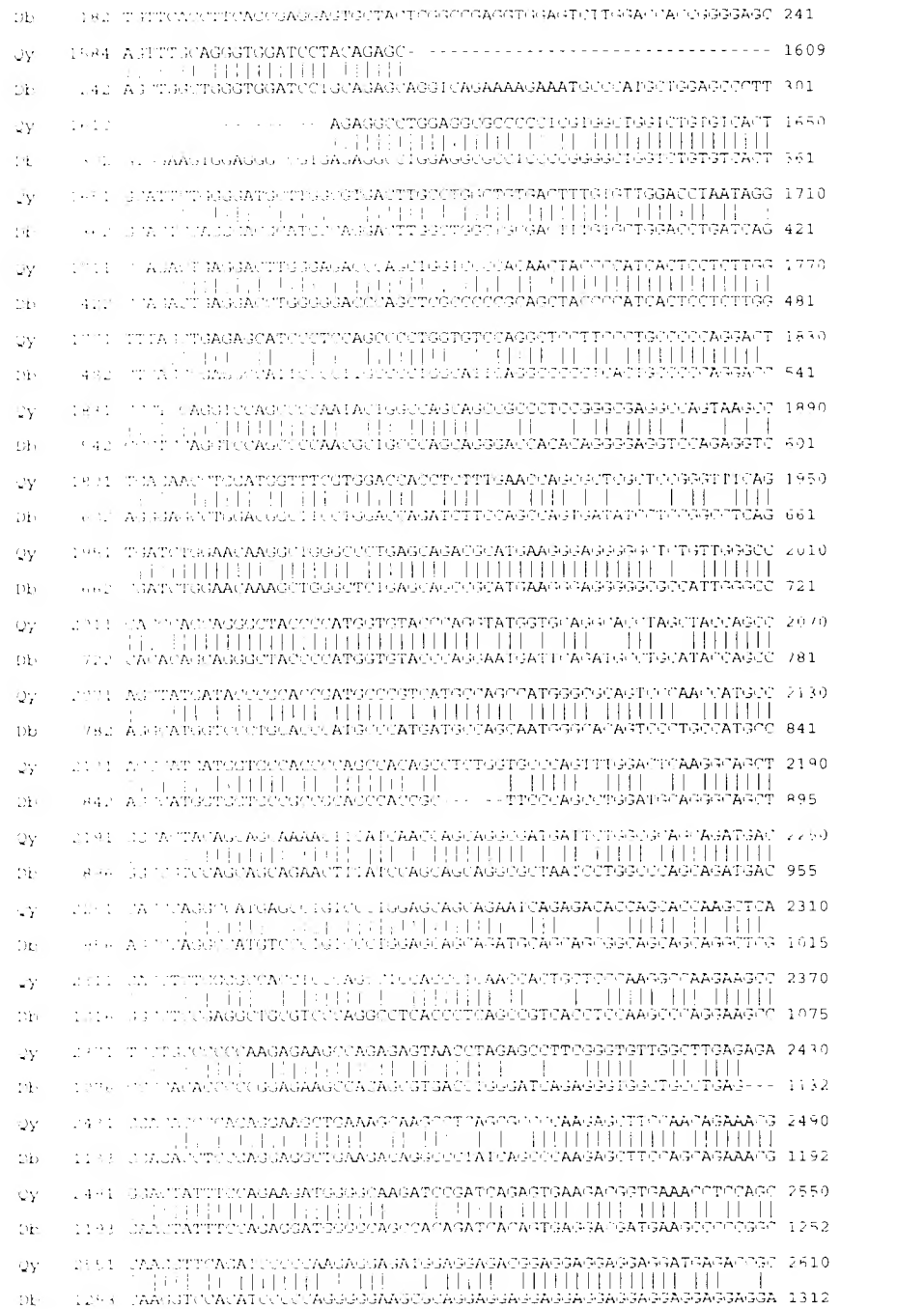
```
b    3961 GGAACCAGGCTGGCAGTTCGGTTCTGACCGGGGCCGCTCCGGACTCTTTCCCGATGACGT 4020

y    4021 GGTGCAGCCAGCTGCTGCCCCCGACCTCTCCTTTTCCCTGGGAAAGAGAAACAGCTGGCA 4080
b    4021 GGTGCAGCCAGCTGCTGCCCCCGACCTCTCCTTTTCCCTGGGAAAGAGAAACAGCTGGCA 4080

y    4081 ACGCAAGAGTAAGCTGGGGCCAGCTCAGGAGGTGAGGAAGACAGAAGAGGTGAAGTGATA 4140
b    4081 ACGCAAGAGTAAGCTGGGGCCAGCTCAGGAGGTGAGGAAGACAGAAGAGGTGAAGTGATA 4140

y    4141 CAGGCCTAACTTGGAGACTGAGAAGGAAAGAGCAGGGTTGCTTCGGGTGTTGTCCACTTC 4200
b    4141 CAGGCCTAACTTGGAGACTGAGAAGGAAAGAGCAGGGTTGCTTCGGGTGTTGTCCACTTC 4200

y    4201 CTGTCCTGGTGGCCAGGGCTCAATGTGTTCCTGTCCTTTACCATCTCCTGACTTTTTGCC 4260
b    4201 CTGTCCTGGTGGCCAGGGCTCAATGTGTTCCTGTCCTTTACCATCTCCTGACTTTTTGCC 4260

y    4261 ATTTGTGAGACTGTAAGTCACACCCCTAACTCTGGTACTTAGTTCAGTGTCTCCATAGA 4320
b    4261 ATTTGTGAGACTGTAAGTCACACCCCTAACTCTGGTACTTAGTTCAGTGTCTCCATAGA 4320

y    4321 GGATGCTTAATAAATAACCTTGGTTTTCCTGG 4352
b    4321 GGATGCTTAATAAATAACCTTGGTTTTCCTGG 4352


SULT 3
058686
CUS       AB058686                6004 bp    mRNA    linear   PRI 02 JUN 2001
FINITION  Homo sapiens mRNA for KIAA1783 protein, partial cds.
CESSION   AB058686
RSION     AB058686.1  GI:14017782
YWORDS    .
URCE      Homo sapiens brain cDNA to mRNA, clone.fg01285.
 ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FERENCE   1  (sites)
 AUTHORS  Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.
 TITLE    Prediction of the coding sequences of unidentified human genes. XX.
          The complete sequences of 100 new cDNA clones from brain which code
          for large proteins in vitro
 JOURNAL  DNA Res. 8 (2), 85-95 (2001)
 MEDLINE  21245130
FERENCE   2  (bases 1 to 6004)
 AUTHORS  Ohara,O., Nagase,T. and Kikuno,R.
 TITLE    Direct Submission
 JOURNAL  Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Rsearch Institute,
          Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
          292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
          URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
          Fax:81-438-52-3914)
ATURES             Location/Qualifiers
   source          1..6004
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="fg01285"
                   /tissue_type="brain"
                   /note="vector:pBluescript II SK plus"
   gene            1..6004
                   /gene="KIAA1783"
   CDS             <1081..5763
                   /gene="KIAA1783"
                   /note="Start codon is not identified."
                   /codon_start=1
                   /product="KIAA1783 protein"
                   /protein_id="BAB47412.1"
                   /db_xref="GI:14017783"
                   /translation="PVCCAGSERLDPVGYTASALGITHPCALARMKEASGVGGHSPM
                   CPPHCHMPPGPAGEWPGATVQQPRQRAPTALLQPDAAGPGGGVSAGVAVLGACPSAS
                   EGVLPRPPGRSAPQPPEYPGRPDMAVPGWSAMVPSPLTATSTSQATDHTHLQKSHYHH
                   GDHPSYAKPPLPLPVFTVRHYAGTVTYQVHKFLNPNRDQLDPAVVEMLGQSQLQLVGS
                   LFQEAEPQSPGGPGPPTLASRFQQALEDLIAPLGPSHVYFIQCLTPNPGKLPGLFDVG
                   HVTEQLHQAAILEAVGTRSANFPVRVPFEAFLASFQALGSEGQEDLSDREKCGAVLSQ
                   VLGAESPLYHLGATKVLLQEQGWQRLEELRDQQRSQALVDLHRSFHTCISRQRVLPRM
                   QARMRGFQAPKPYLPRPAALGQLNTILLVAQPLLQRRQRLQLGRWQGWHSSERALERV
                   PSMELGRLEIPAELAVMLKTAESHRDALAGSITECLPPEVPARPSLTLPADIDLFPRS
                   SFVAIGFQEPSLPRPGQPLAKPLTQLDGDNPQRALDINKVMLRLLGDGSLESWQRQIM
                   GAYLVRQGQCWPGLRNELFSQLVAQLWQNPDEQQSQRGWALMAVLLSAFPPLPVLQKP
                   LLKFVSDQAPRGMAALCQHKLLGALEQSQLASGATRAHPPTQLEWLAGWRRGRMALDV
                   FTFSEECYSAEVESWTTGEQLAGWILQSRGLEAPPRGWSVSLHSRDAWQDLAGCDFVL
                   DLISQTEDLGDPARPRSYPITPLGSAEAIPLAPGIQAPSLPPGPPPGPAPTLPSRDHT
                   GEVQRSGSLDGFLDQIFQPVISSGLSDLEQSWALSSRMKGGGAIGPTQQGYPMVYPGM
                   IQMPAYQPGMVPAPMPMMPAMGTVPAMPAMVVPPQPPLPSLDAGQLAVQQQNFIQQQA
                   LILAQQMIAQAMSLSLEQQMQQRQQQARASEAASQASPSAVISKPRKPPTPPEKPQRD
                   LGSEGGCLRETSEEAEDRPYQPKSFQQKRNYFQRMGQPQITVRTMKPPAKVHIPQGEA
                   QEEEEEEEEEEQEEQEVETRAVPSPPPPPIVKKPLKQGGAKAPKEAEAEPAKETAAK
                   GHGQGPAQGRGIVVRSSLSKPPRRPQPSREIGNIIRMYQSRPGPVPVPVQPSRPPKAPL
                   RKIDPKDEALAKLGINGAHSSPPMLSPSPGKSPPPAVAPRPKAPLQLGPSSSIKEKQG
                   PLLDLFGQKLPIAHTPPPPPAPPLPLPEDPGTLSAERRCLTQPVEDQGVSTQLLAPSG
                   SVCFSYTGTPWKLFLRKEVFYPRENFSHPYYLRLLCEQILRDTFSESCIRISQNERRK
                   MKDLLGGLEVDLDSLTTTEDSVKKRIVVAARDNWANYFSRFFPVSGESGSDVQLLAVS
                   HRGLRLLKVTQGPGLRPDQLKILCSYSFAEVLGVECRGGSTLELSLKSEQLVLHTARA
                   RAIEALVELFLNELKKDSGYVIALRSYITDNCSLLSFHRGDLIKLLPVATLEPGWQFG
                   SAGGRSGLFPADIVQPAAAPDFSFSKEQRSGWHKGQLSNGEPGLARWDRASERRKMGE
                   GQAEARPA"
BASE COUNT     1134 a   1942 c   1870 g   1058 t
ORIGIN

  Query Match             27.9%;  Score 2284.2;  DB 9;  Length 6004;
  Best Local Similarity   75.7%;  Pred. No. 0;
  Matches 3397;  Conservative    0;  Mismatches  928;  Indels   66;  Gaps    9;

Qy          2 GCTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGAC 61
Db       1624 GCAGGGACTGTCACCTACCAGGTTCACAAGTTTTAAACAGAAACCGGGATCAGCTGGAC 1683

Qy         62 CCCGCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCAG 121
Db       1684 CCTGCTGTGGTGGAGATGCTTGGCCAGAGCCAGC-------------------TCCAG 1722

Qy        122 CTCATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGGCTGGGACTCACCAAAACAAACCC 181
Db       1723 CTGGTGGGCAGCCTGTTCCAGGAGGCAGAGCCCCAGTCCAGGGGAGGGCGAGGCAGACCC 1782

Qy        182 ACATTGGCCTCTCGATTCCAGCAGACCCTGGGTGACTTGCTAGCTGGGCTAGGCAGCAGC 241
Db       1783 ACGCTGGCCTCTCGCTTCCAGCAGGCCCTGGAGGACTTCATAGCCCGGCTGG---GCAGG 1839

Qy        242 GGCCATGTCTACGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATGCCAGGCCTCTTC 301
Db       1840 AGCCATGTCTATTTCATCCAGTGCCTCACCCCTAACCCTGGAAAGCTTCCAGGCCTCTTT 1899

Qy        302 GACGTGGGCCATGTGGCAGAGCAGCTGCGTCAGGCTGGCATCCTGGAGATCATAGGCACC 361
Db       1900 GACGTGGGACATGTGACAGAGCAACTGCACCAGGCAGCCATACTGGAGGCCGTGGGCACC 1959

Qy        362 CGGAGTACCCACTTCCCCGTGCGGAGCCCCTTCCAAGTCTTCTTGCCTAAGGTTCCATGC 421
Db       1960 CGCAGTGCCAACTTCCCCGTGCGTGTGCCCTTTGAGGCCTTCCTGGCCAGCTTCCAGGCC 2019

Qy        422 CTGGGGTCAGGGAGACAGAAAGCTGGCTCTGACCAGGAGAGGTGTGGTGCCATCCTCAGC 481
Db       2020 CTGGGGTCAGAAGGGCAGGAAGACCTCTCTGACCGGGAGAAGTGTGCTGCCGTCCTGAGC 2079

Qy        482 GAAGTGCTGGGGGTAGAGTCACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCAG 541
Db       2080 CAGGTGCTGGGGGCGGAATCACCTCTCTATCACCTTGGAGCCACCAAGGTCCTGCTGCAG 2139

Qy        542 GAACAGGGCTGGCAGCAGCTAGAACAGCTGTGGGCTCAGGAGCGTCACAGGCCTGCTC 601
Db       2140 GAGCAGGGCTGGCAGCGGCTGGAGGAGCTCCGGGACCAGCAGCGCTCCCAGGCCCTGGTC 2199

Qy        602 ACTCTGCACCGTGGCCTCCGAGCTGTATCACCCGCAGCGCTCCGTCTCCTGCCCCGG 661
```

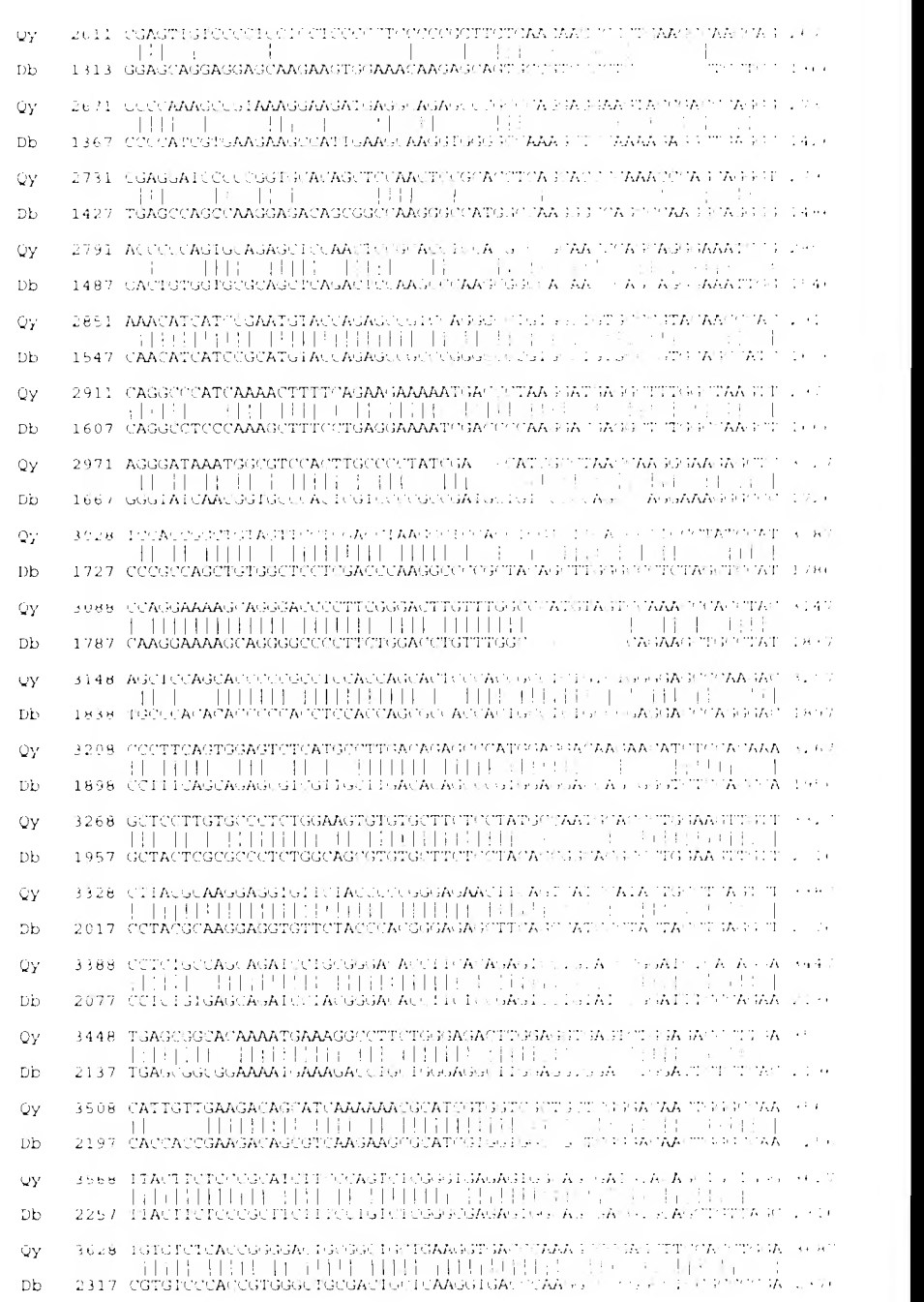


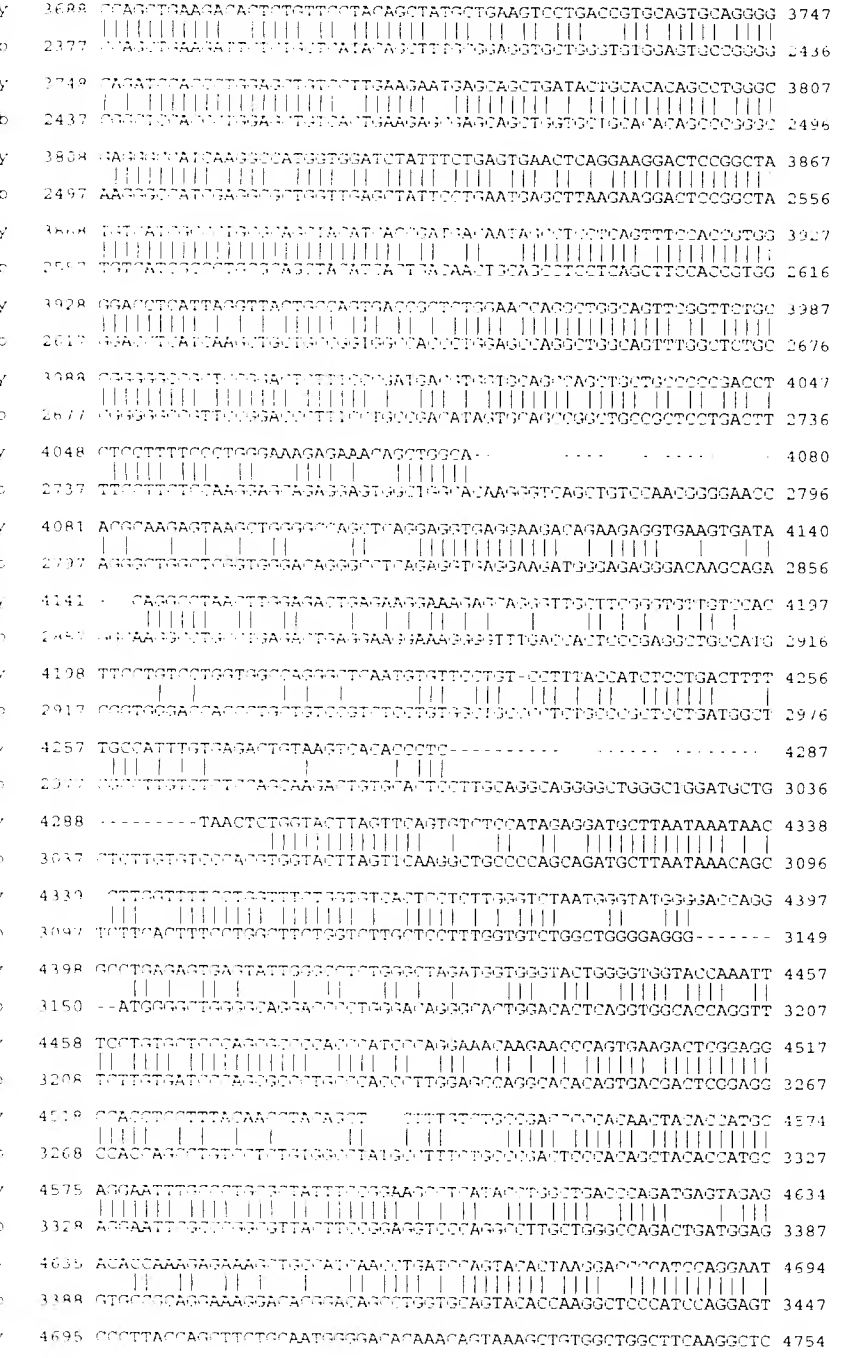


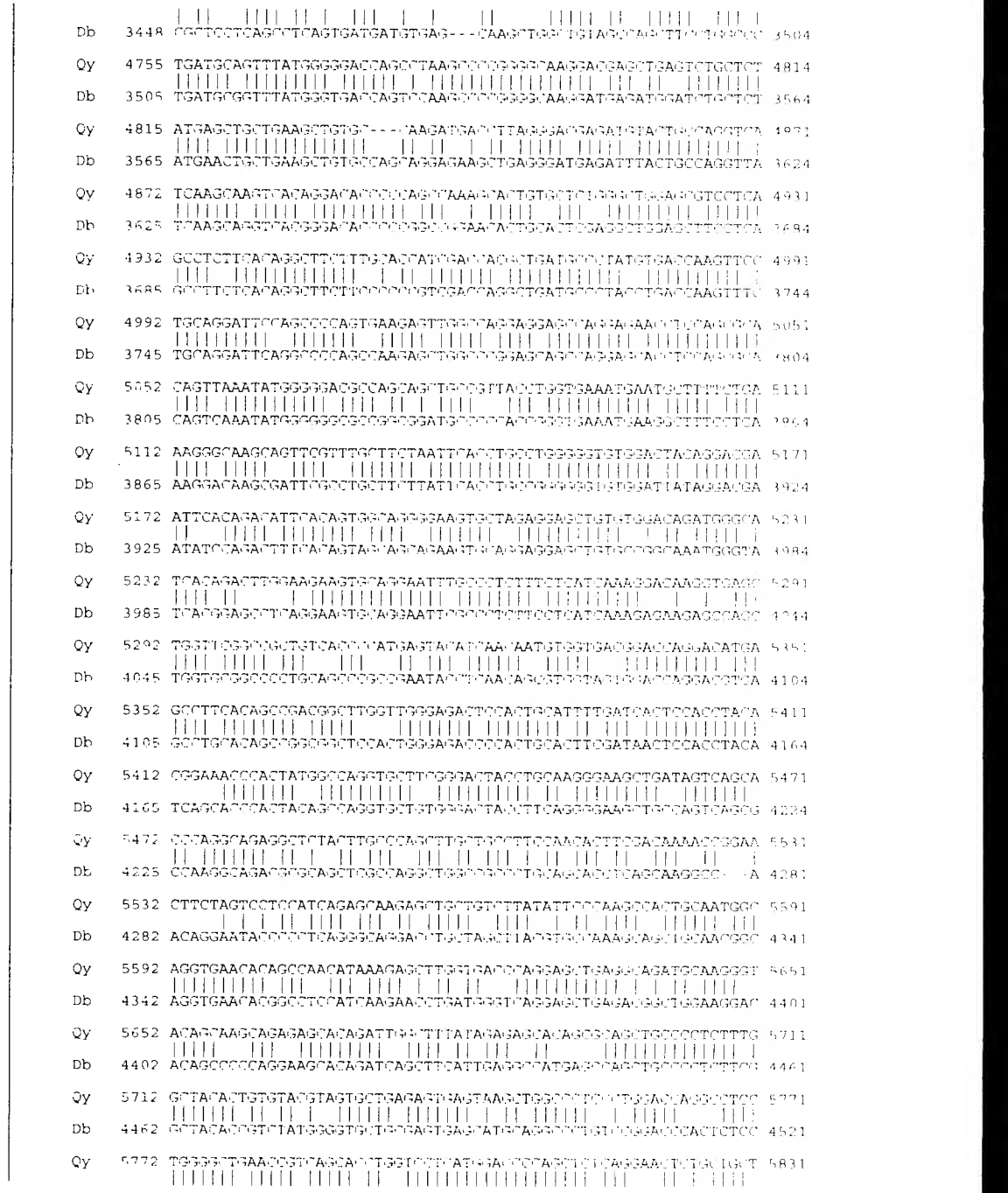
```
Qy   2813 TCCGCACCTCCACGCCCGCAACCCAGCAGGGAAATCCGAAACATCATCCGAATGTACCAG 2872
          |||   ||     || ||  ||||||||||||||||||   | ||||||||||| ||||||||
Db   4393 TCTAAGCCTAAGCGGCCACAACCCAGCAGGGAAATTGGCAACATCATCCGCATGTACCAG 4452

Qy   2873 AGCCGTCCAGGGCCTTGTGCTGTCCCGTACAACCACCAGGCCCATCAAAACTTTTCAG 2932
          ||||| || || || ||| |||||||||| || ||  |||||||   |||| |||| | |
Db   4453 AGCCGCCGGGCCCCCTGCCTGTGCCCGTGCAGCCATCCAGGCTCCCAAAGCTTTCCTG 4512

Qy   2933 AAGAAAAATGACCCTAAGGATGAGGCTTTGGCTAAGTTAGGGATAAATGGCGTCCACTTG 2992
          | |||||| ||||| ||||| |||||| |||| ||| | || || || || | |||||| |
Db   4513 AGGAAAATCGACCCCAAGGACGAGGCTCTGGCCAAGCTGGGTATCAACGGTGCCCACTCG 4572

Qy   2993 CCCTTATCGA---CATCGCCTAACCAAGGGAAGAGCTCTCCACCGGCTGTAGTTCCTCGA 3049
          |||   |||     || || | || ||| ||| || | || || ||||| | |||||||
Db   4573 TCCCCCCCGATGCTGTCCCCCAGCCCAGGAAAGGGCCCCCCGCCAGCTGTGGCTCCTCGA 4632

Qy   3050 CCTAAGGCTCGACTTGGTCTTGAGCCTTCCCTATCCATCCAGGAAAAGCAGGGACCCCTT 3109
          || |||||  |  |  |||| ||| ||     |||||| |||||||||||||| ||||||
Db   4633 CCCAAGCCCCCCTACAGTTTGGCCCCTAGCTCCATCAAGGAAAAGCAGGGGCGCCCTT 4692

Qy   3110 CGGGACTTGTTTGGCCCATGTAGTCCAAACCCACCTACAGCTCCAGCACCCCCGCCTCCA 3169
          | ||||  ||||||||          || |  ||||  ||  |  |||||| ||||||
Db   4693 CTGGACCTCTTTGGCCAG---------AAGCTGCCTATTGCCCACACACCCCCACCTCCA 4743

Qy   3170 CCAGCACTCCCACCCCCCTCTCCTCCCGAGCCCAAGACCCCTTCAGTGGAGTCTCATGCC 3229
          |||||  ||||  |||||| | | ||| ||   ||||| |||||   |||  ||  |  |
Db   4744 CCAGCCCCACCACTGCCTGTCCCCGAGGACCCAGGGACCCTTTCAGCAGAGCGCTGCTTGC 4803

Qy   3230 TTGACAGAGCCCATGGAGGACAAGAACATCTCCACAAAGCTCCTTGTGCCCTCTGGAAGT 3289
          |||||| |||||  |||||||| ||     |||||||  |||| || | |||||||| ||
Db   4804 TTGACACAGCCCGTGGAGGACCAGGGGGTCTCCACCCAGCTACTCGCCGCCCTCTGGCAGC 4863

Qy   3290 GTGTGCTTCTCCTATGCCAATGCACCCTGGAAGTTGTTCTTACCCAAGGAGGTGTTCTAC 3349
          ||||||||||||||  ||    | ||||||||||||| ||||||||||||||||||||||
Db   4864 GTGTGCTTCTCCTACACCGGCAGCCCCTGGAAGTTGTTCTTACCCAAGGAGGTGTTCTAC 4923

Qy   3350 CCCCGGGAGAACTTCAGTCATCCTATCATCCTGAGGCTCCTCCTGCCAGCAGATCCTGCGG 3409
          || |||||||||||||| |||| |||| ||| || ||||||||  ||||||||| |||
Db   4924 CCACGGGAGAACTTCAGCCATCCCTACATCCTGAGGCTCCTCCTGCCAGCAGATCCTACGG 4983

Qy   3410 GACACCTTCACAGAGTCCTCCACCCCGAGCCTCACAGGATGAGCGGCACAAAATGAAAGGC 3469
          |||||||| | |||||||| | |||||| || ||| |||||||||  |||||||||| |
Db   4984 GACACCTTCTCCGAGTCCTCTATCCCGATTTCCCAGAATGAGCCGGCCGAAAATGAAAGAC 5043

Qy   3470 CTTCTGGGAGACTTGGAGTGAGTCCTGAACACCCTTGACATTGTTGAAGACAGCATCAAA 3529
          || ||||||| |||||||||   |||||  | ||   ||          |||||||  ||||
Db   5044 CTGCTGGGAGGCTTGGAGTGGGACCTGGATTGTCTCACCACCCGAAGACAGCCTCAAG 5103

Qy   3530 AAACCCATCCCTGGTCGCTGCTCGGGACAACTGGGCCAATTACTTCTCCCCGCATCTTCCCA 3589
          || |||||||||| || ||||||||||||||||||||||||||||||||||| |||| ||
Db   5104 AAGCCCATCGTGGTGCCCCGCTCGGGACAACTGGCCCAATTACTTCTCCCGCTTCTTTCCT 5163

Qy   3590 GTCTCGGGTGAGAGTGGCAGCTATGTACAGCTGCTGGGTGTGTCTCACCGGGGACTGCGG 3649
          |||||||| ||||||||||||| || |||||  | |  |||||  |||||  || |||||
Db   5164 GTCTCGGGCCAGAGTGCCAGCGAGTCCACCTGTTAGCCGTGTCCCACCGTGGGCTGCGA 5223

Qy   3650 CTGCTGAAGGTGACCCAAAGCCCGAGCTTCCACCTGGACCAGCTGAAGACACTCTGTTCC 3709
          ||||| |||||||||| ||||   || || ||  |||||||||||||  ||||| |||||
Db   5224 CTGCTCAAGGTGACCCAAGGCCCCGGCCTCCGCCCCGACCAGCTGAAGATTCTCTGCTCA 5283

Qy   3710 TACAGCTATGCTGAAGTCCTGACCGTGCAGTGCAGGGGCAGATCCACCCTGGAGCTGTCC 3769
          ||||||| ||| || || |||   ||| ||||| ||||| | ||||||||||||||||||
Db   5284 TACAGCTTTGCCGAGGTGCTGGGTGTGGAGTGCCGGGGCGGCTCCACCCTGGAGCTGTCA 5343

Qy   3770 TTGAAGAATGAGTACCTGATACCTGACACAGCCTGGGCGAGGGCCATCAAGGCCATGGTG 3829
          |||||| |||||||| | |||||||||||| |||| |||||||| ||||  ||||
Db   5344 CTGAAGAGTGAGTACCTGGTGCTGCACACAGCCCGGGCAAGGGCCATCGAGGCGCTGGTT 5403

Qy   3830 GATCTATTTCTGAGTGAACTCAGGAAGGACTCCGGCTATGTCATCGCCCTGCGCAGCTAC 3889
          || ||||| |||| || | |||||||||||||||||||||||||||||||||||||||||
Db   5404 GAGCTATTCCTGAATGAGCTTAAGAAGGACTCCGGCTATGTCATCGCCCTGCGCAGCTAC 5463

Qy   3890 ATCACCGATGACAATAGCCTCCTCAGTTTCCAGCGTGGGGACCTCATTAGGTTACTGCCA 3949
```

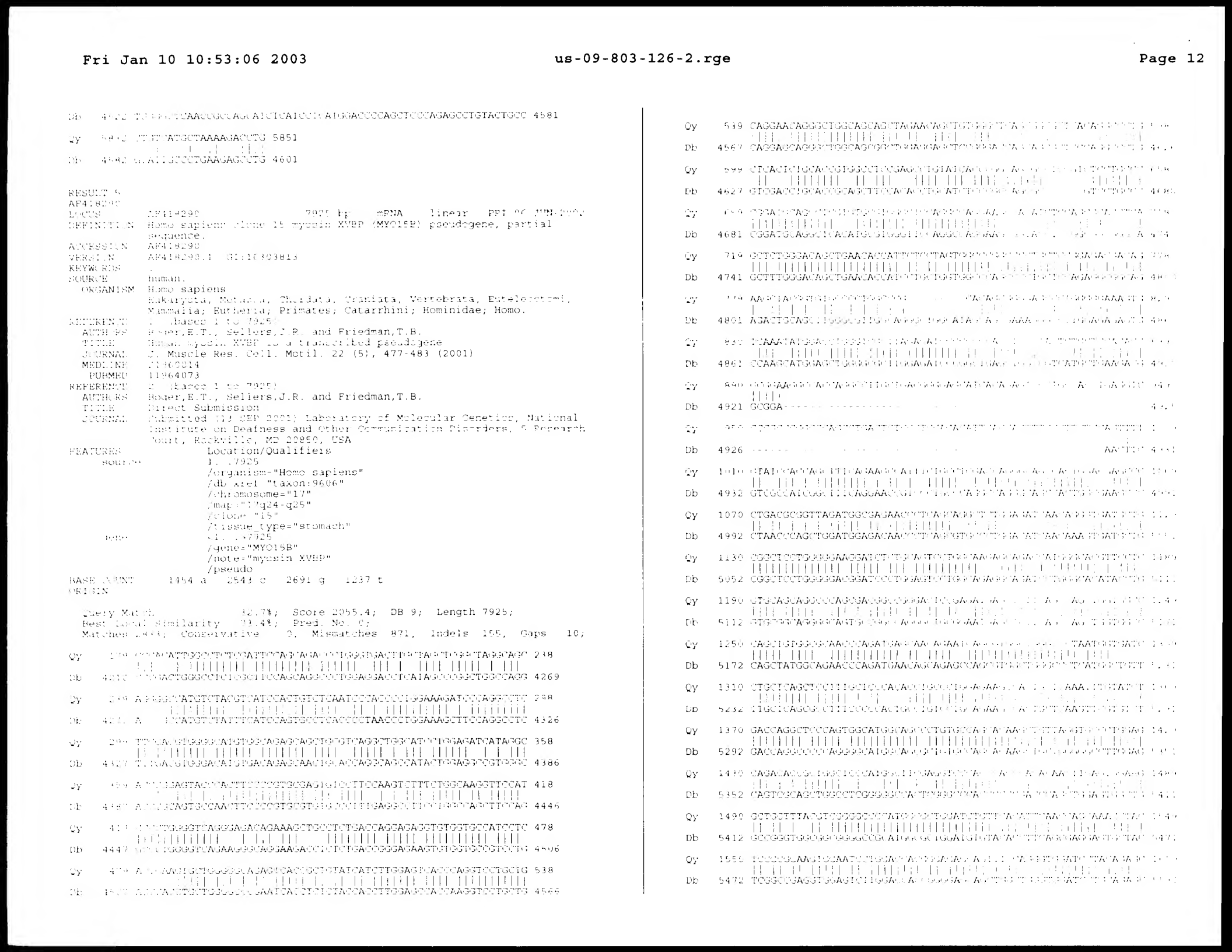
```
          |||||  ||   ||   |||||||||| ||||||||||||||||||| | | | |||||
Db   5464 ATCACTGACAACTGCAGCCTCCTCAGCTTCCACCGTGGGGACCTCATCAAGCTGCTGCCC 5523

Qy   3950 GTGACCGCTCTGGAACCAGGCTGGCAGTTCGGTTCTGCCGGGGGCCGCTCGGACTCTTT 4009
          ||| || | ||||| |||||||||||| || |||||||||||||| ||||||||||||
Db   5524 GTGGCCACCCTGGAGCCAGGCTGGCAGTTTGGCTCTGCCGGGGGCCGTTCCGGACTCTTT 5583

Qy   4010 CCCGATGACGTGGTGCAGCCAGCTGCTGCCCCCGAGCTCTCTTTCCCTGGGAAAGAGA 4069
          || |  ||| | |||||||| ||||| || |||||| | |||| | |||  ||  ||||
Db   5584 CCTGCCGACATAGTGCAGCCGGCTGCAGCTCCCGACTTTCCTTCTCCAAGGAGCAGAGG 5643

Qy   4070 AACAGCTGGCA 4080
          |   |||||||
Db   5644 AGTGGCTGGCA 5654


RESULT 4
AF418292
LOCUS       AF418292                4602 bp    cDNA    linear   PRI 16-JUN-2002
DEFINITION  Homo sapiens clone cDNA1 myosin XVBP (MYO15B) pseudogene, partial
            sequence.
ACCESSION   AF418292
VERSION     AF418292.1  GI:16303815
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4602)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Human myosin XVBP is a transcribed pseudogene
  JOURNAL   J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
  MEDLINE   21960014
   PUBMED   11964073
REFERENCE   2  (bases 1 to 4602)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2001) Laboratory of Molecular Genetics, National
            Institute on Deafness and Other Communication Disorders, 5 Research
            Court, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .4602
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q24-q25"
                     /clone="cDNA1"
                     /tissue_type="kidney; stomach"
     gene            <1. .>4602
                     /gene="MYO15B"
                     /note="myosin XVBP"
                     /pseudo
BASE COUNT      927 a   1495 c   1405 g    775 t
ORIGIN

  Query Match           35.7%;  Score 2246;  DB 9;  Length 4602;
  Best Local Similarity   71.0%;  Pred. No. 0;
  Matches 3293;  Conservative    0;  Mismatches 1175;  Indels  172;  Gaps   17;

Qy   1344 AGAAGCCACTGCTCAAATTGGIAICIGACCAGGCTCCCAGTGGCATGGCAGCCCTGTGCC 1403
          |||||||||||||||| || || |||||||||| ||||| ||||||||| |||||||
Db      2 AGAAGCCACTGCTCAAGTTCGTGTCTGACCAGGCCCCCAGGGCCATGGCAGCCCTGTGCC 61

Qy   1404 AGCACAAGCTGTTAGGTGGCCTGGAGCAGACACCGCTGGCTCCCATGGCTTCGAGGTCCC 1463
          |||||||||| | || ||| |||||||| | | ||||||   |||  |  || |||  |||
Db     62 AGCACAAGCTGCTGGGGGCCTTGGAGCAGTGGCAGCTGGCCTGGGGGCCACCGGGCCC 121

Qy   1464 ACCCACCCACACAACTTGAGTGGAAGGCTGGTTTACGTCGGGGCCGCATGGCGCTGGATG 1523
          |||| || || || || |||||||  ||| || |  || ||||||||||||||||||||||
Db    122 ACCCCCCGACCCAGCTGGAGTGGCTGGCCGGATGGCGGCGGGGCCGCATGGCGCTGGATG 181

Qy   1524 TGTTCACATTCAAGGAGGAAAGCTACTCCGCGGAAGTGGAATCCTGGACCAGGGAGAGC 1583
          ||||||| |||| ||||||  ||||||| || || ||||| || |||||||| || ||||
```



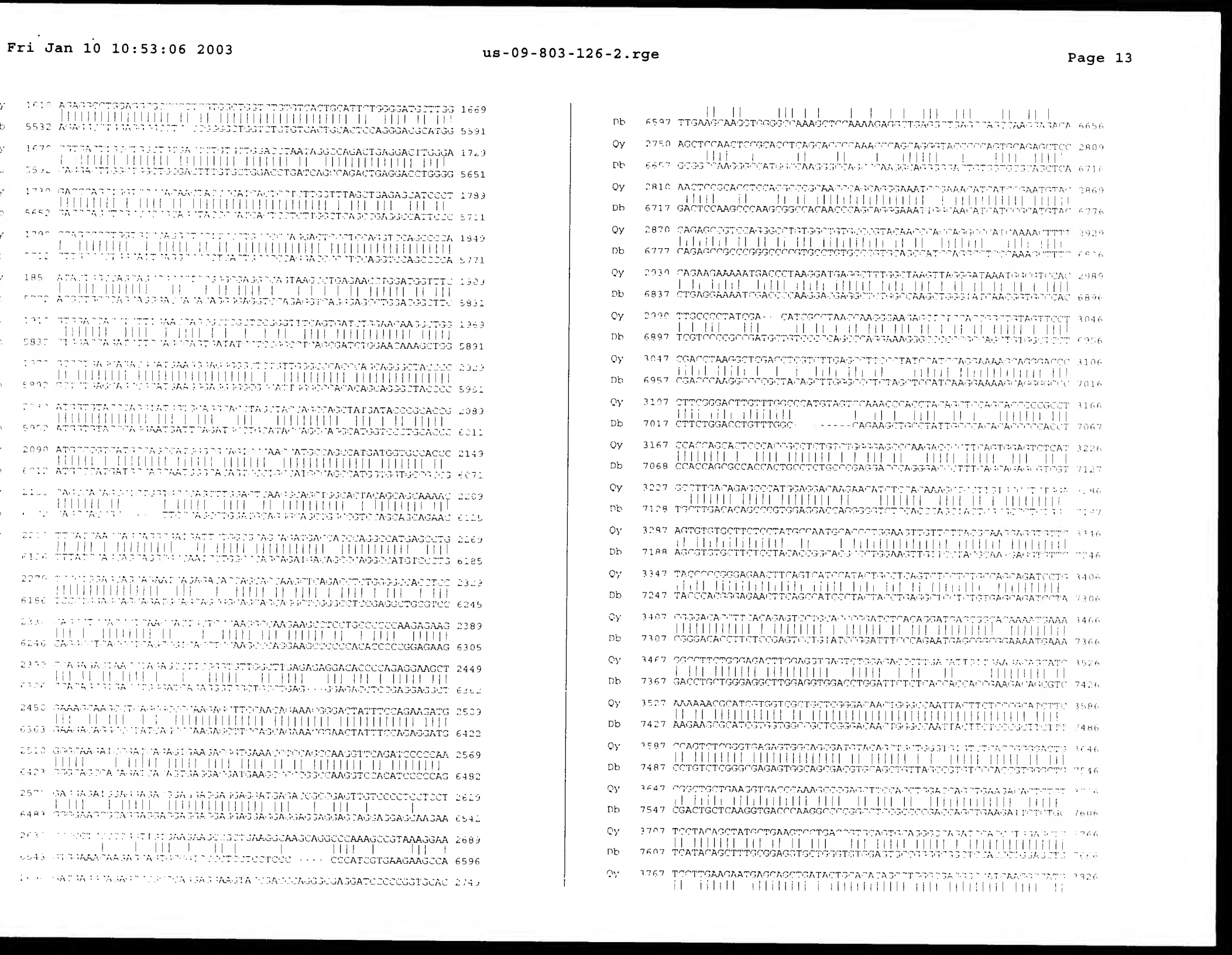


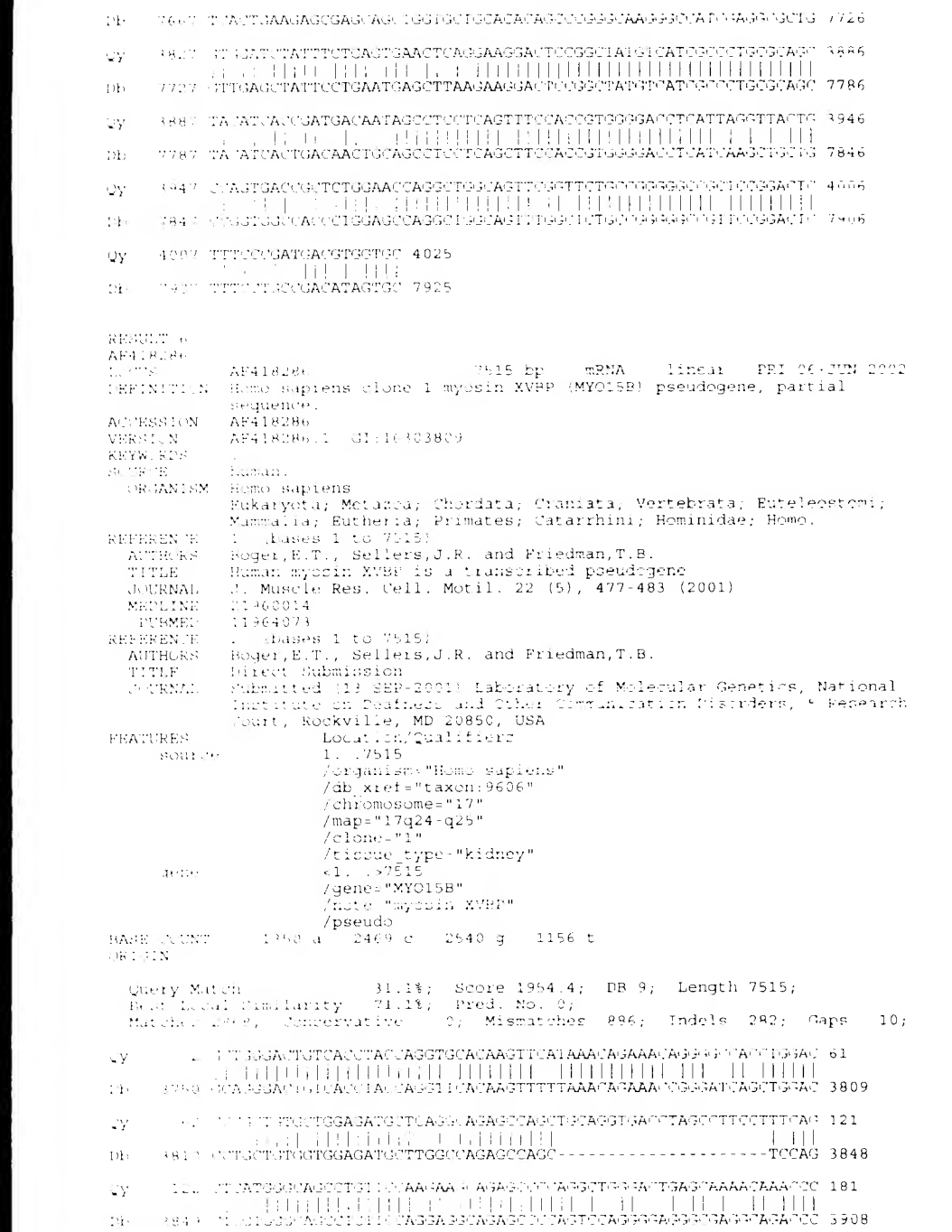




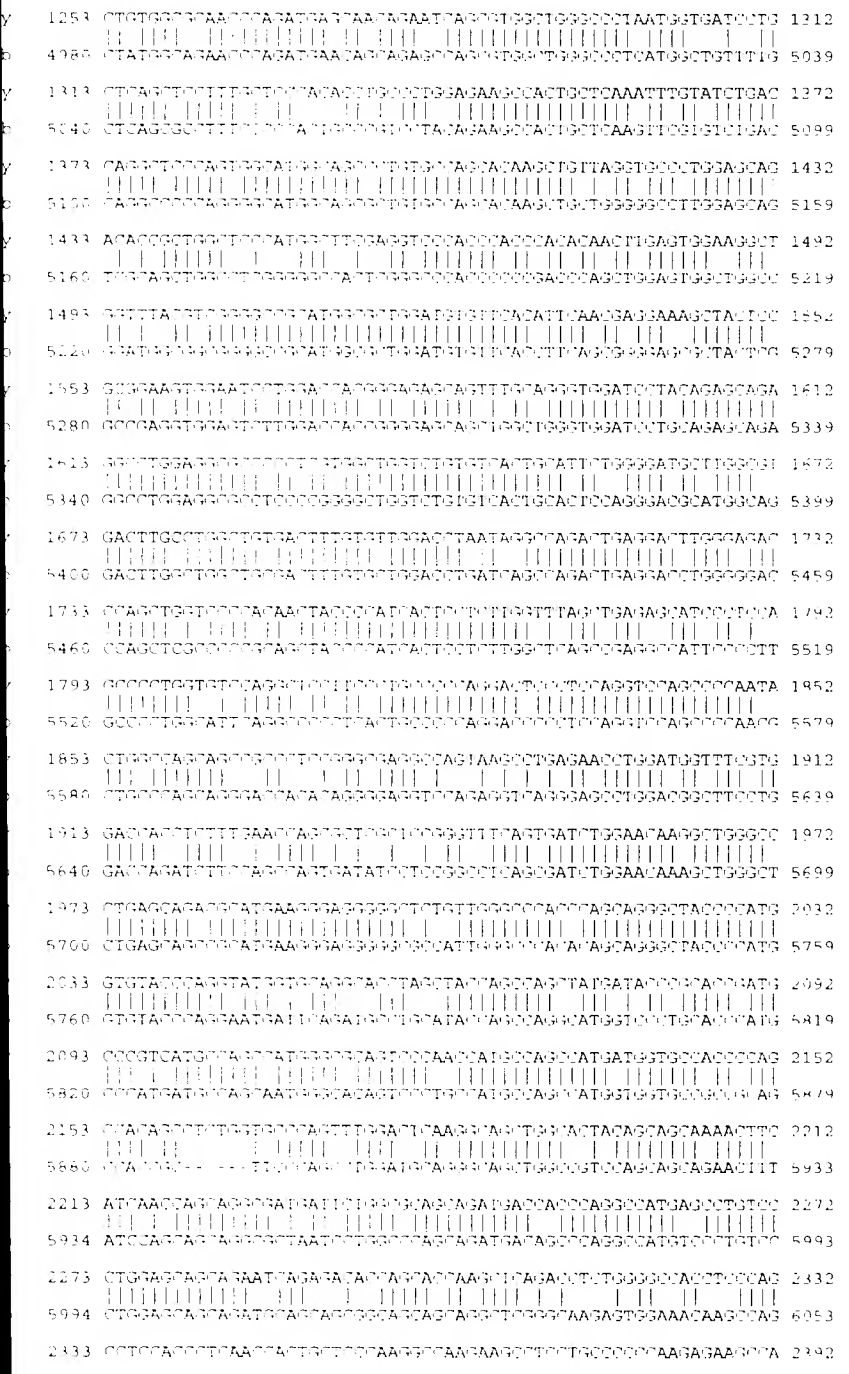


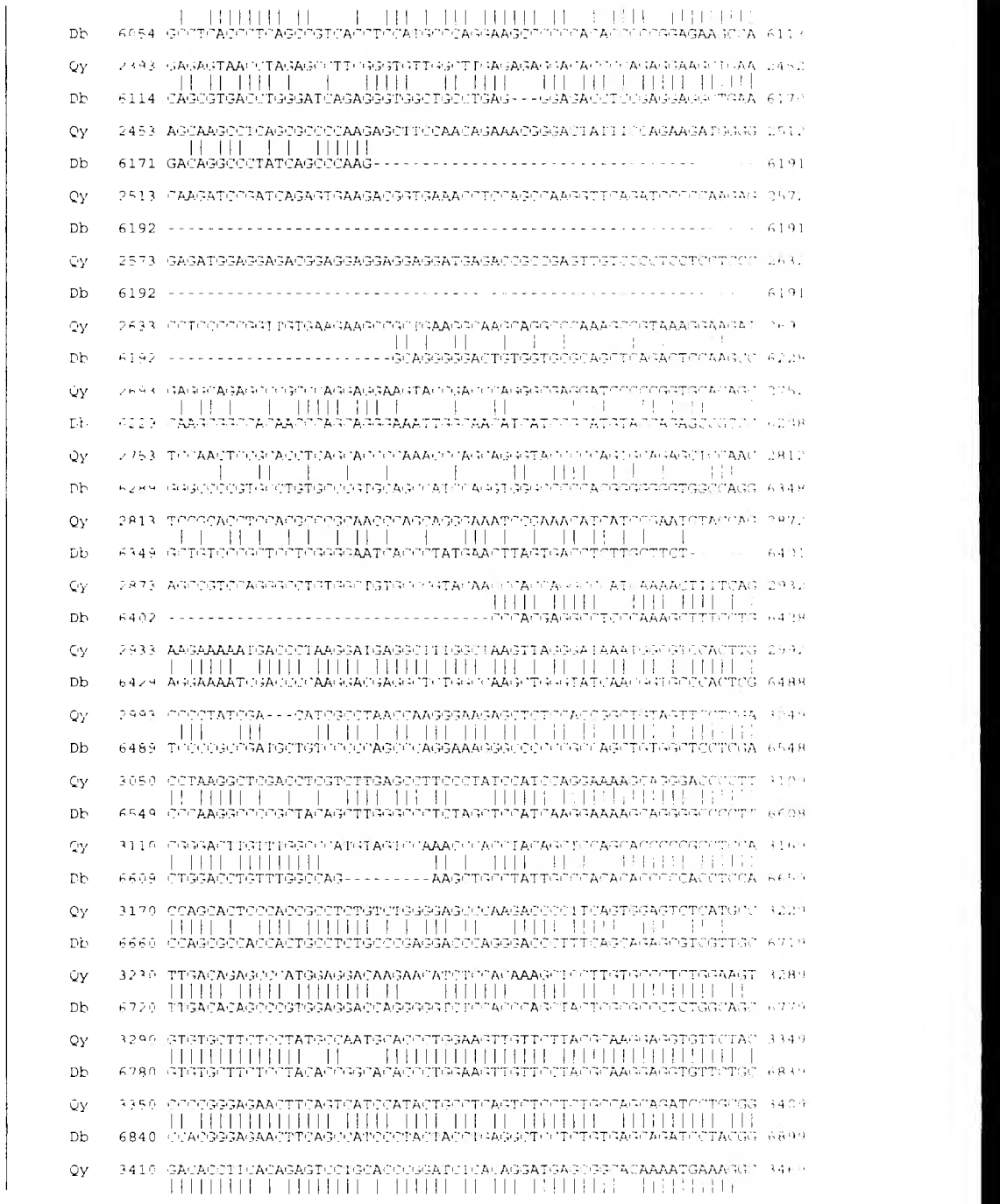
```
RESULT 6
AF418290
LOCUS       AF418290                7925 bp    mRNA    linear   PRI 06-JUN-2001
DEFINITION  Homo sapiens clone 15 myosin XVBP (MYO15B) pseudogene, partial
            sequence.
ACCESSION   AF418290
VERSION     AF418290.1  GI:16303814
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7925)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Human myosin XVBP is a truncated pseudogene
  JOURNAL   J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
  MEDLINE   21960014
    PUBMED  11964073
REFERENCE   2  (bases 1 to 7925)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2001) Laboratory of Molecular Genetics, National
            Institute on Deafness and Other Communication Disorders, 5 Research
            Court, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..7925
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q24-q25"
                     /clone="15"
                     /tissue_type="stomach"
     gene            <1..>7925
                     /gene="MYO15B"
                     /note="myosin XVBP"
                     /pseudo
BASE COUNT     1454 a   2543 c   2691 g   1237 t
ORIGIN

  Query Match          82.7%;  Score 2055.4;  DB 9;  Length 7925;
  Best Local Similarity   93.4%;  Pred. No. 0;
  Matches 1883;  Conservative    0;  Mismatches  871;  Indels  155;  Gaps   10;
```

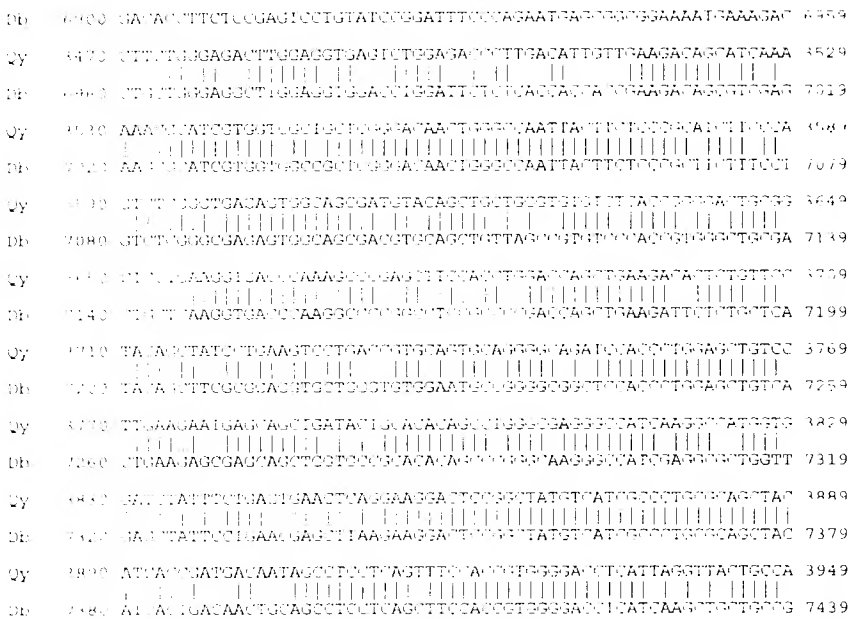




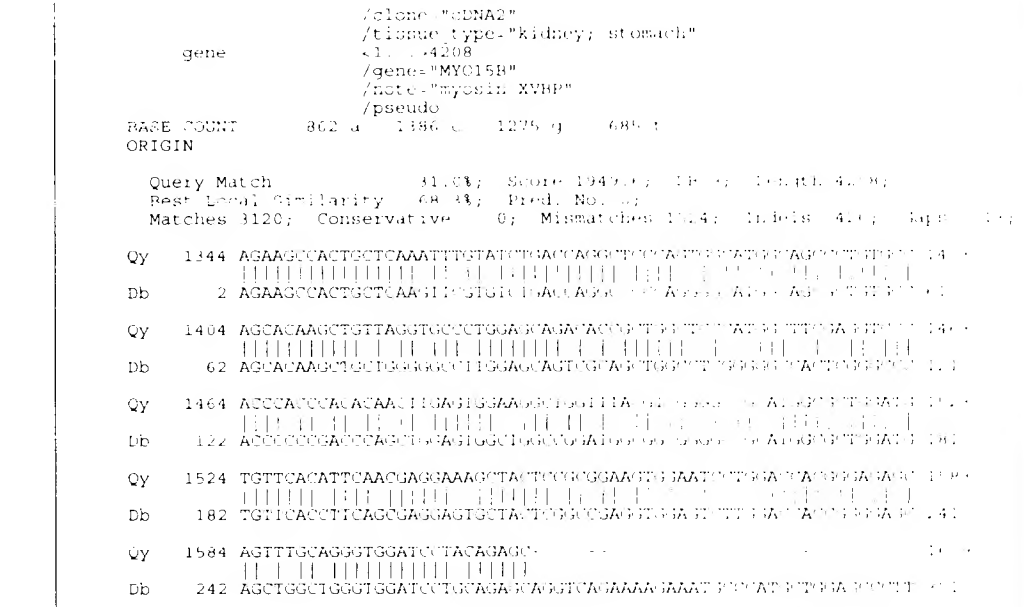
```
RESULT 6
AF418286
LOCUS       AF418286               7515 bp    mRNA    linear   PRI 06-JUN-2002
DEFINITION  Homo sapiens clone 1 myosin XVBP (MYO15B) pseudogene, partial
            sequence.
ACCESSION   AF418286
VERSION     AF418286.1  GI:16303809
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7515)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Human myosin XVBP is a transcribed pseudogene
  JOURNAL   J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
  MEDLINE   21960014
    PUBMED  11964073
REFERENCE   2  (bases 1 to 7515)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2001) Laboratory of Molecular Genetics, National
            Institute on Deafness and Other Communication Disorders, 5 Research
            Court, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..7515
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q24-q25"
                     /clone="1"
                     /tissue_type="kidney"
     misc_feature    <1..>7515
                     /gene="MYO15B"
                     /note="myosin XVBP"
                     /pseudo
BASE COUNT     1950 a   2469 c   2540 g   1156 t
ORIGIN

  Query Match             31.1%;  Score 1954.4;  DB 9;  Length 7515;
  Best Local Similarity   71.1%;  Pred. No. 0;
  Matches 2249;  Conservative    0;  Mismatches  886;  Indels  282;  Gaps   10;
```

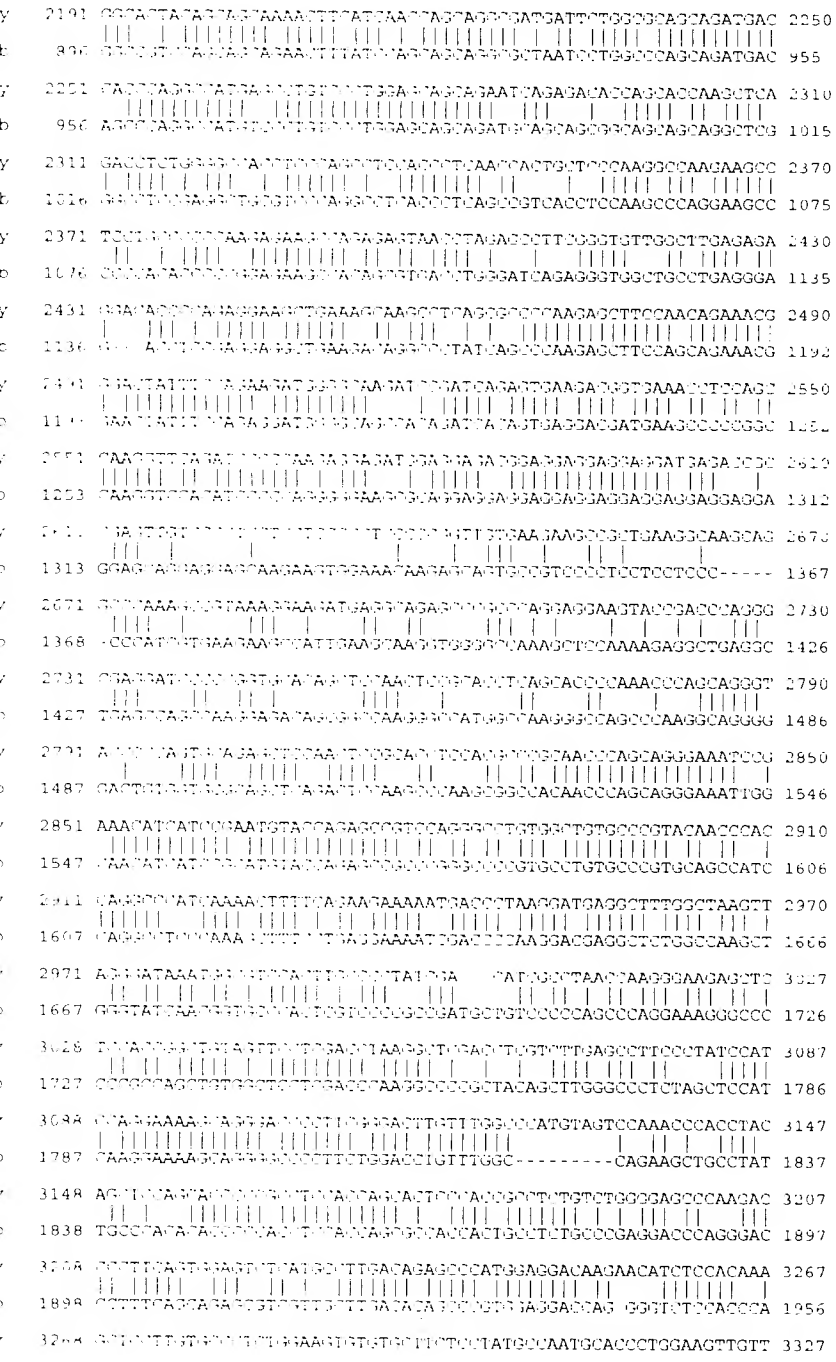


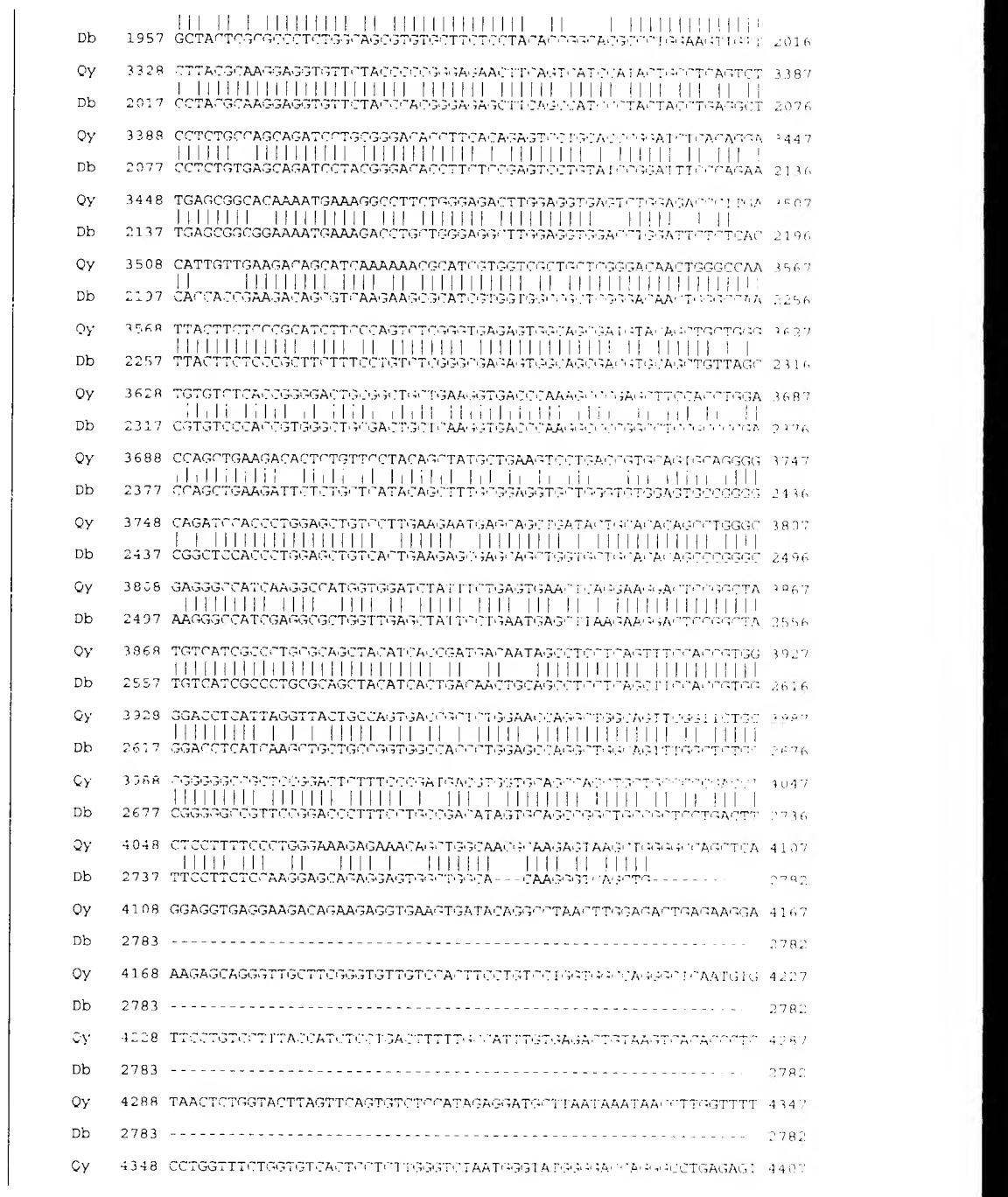


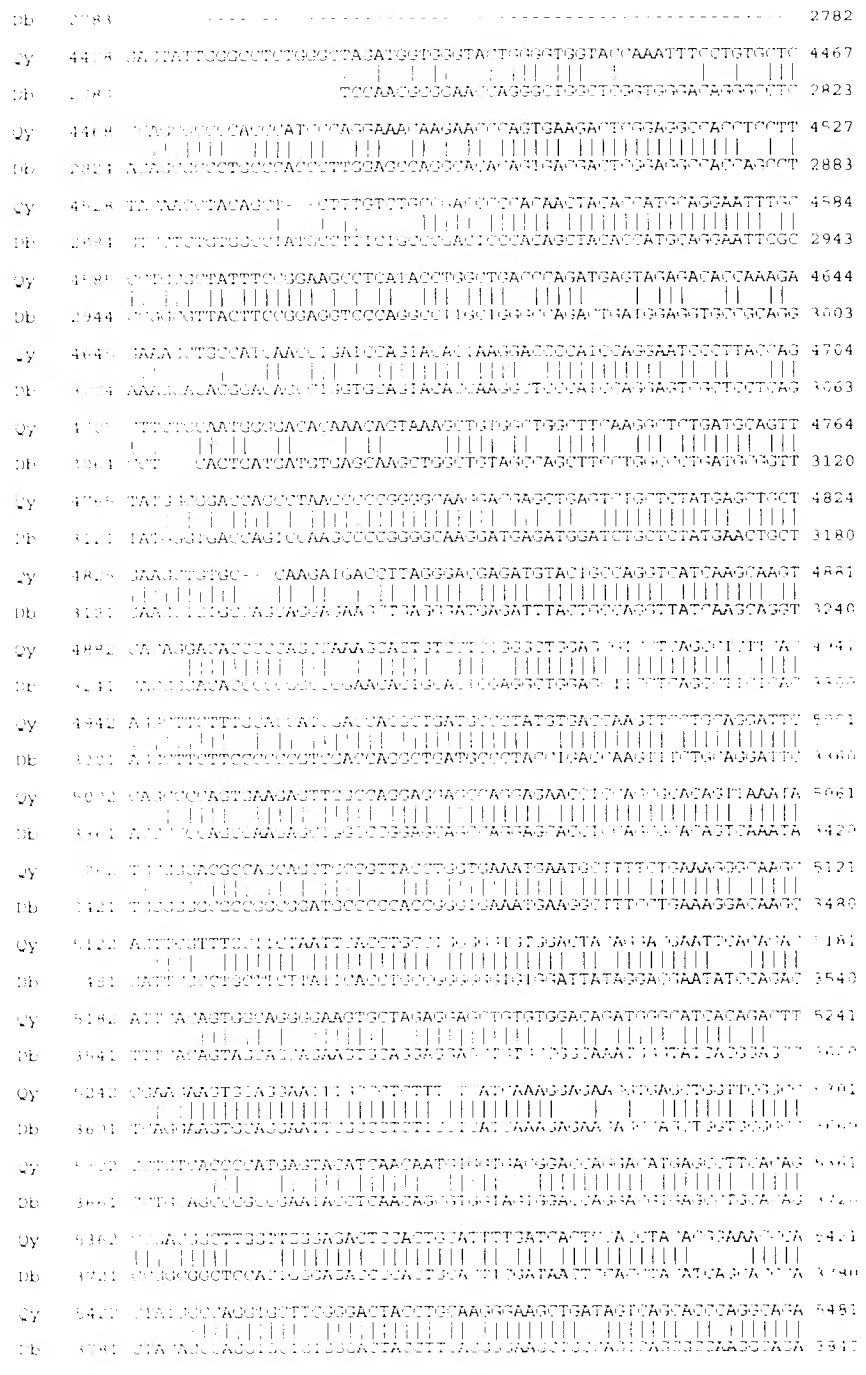


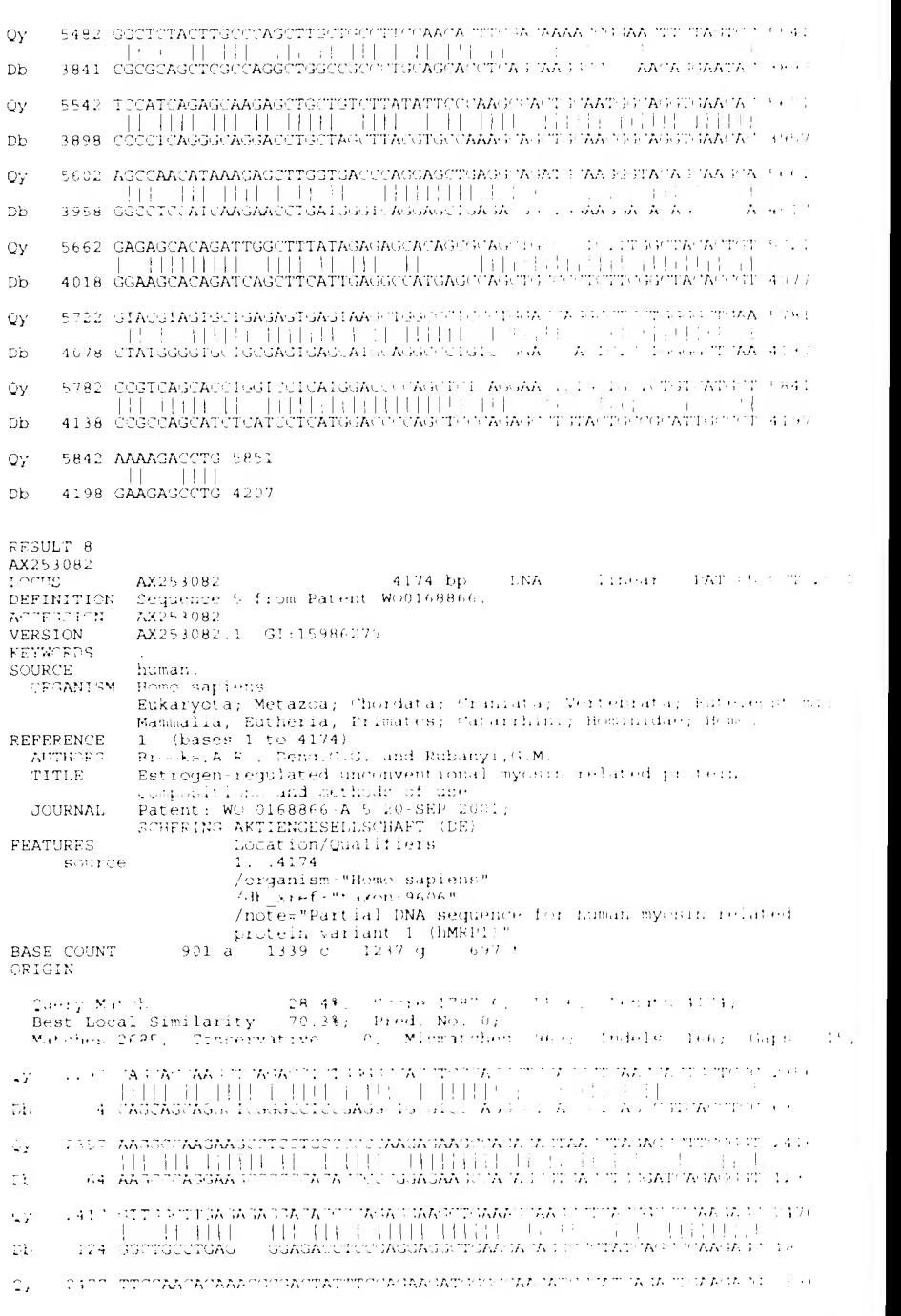
RESULT 2
AF418293
LOCUS AF418293 4208 bp mRNA linear PRI 07-JUN-2001
DEFINITION Homo sapiens clone cDNA2 myosin XVBP (MYO15B) pseudogene, partial sequence.
ACCESSION AF418293
VERSION AF418293.1 GI:16303816
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4208)
AUTHORS Boger,E.T., Sellers,J.R. and Friedman,T.B.
TITLE Human myosin XVBP is a transcribed pseudogene
JOURNAL J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
MEDLINE 21760014
PUBMED 11964073
REFERENCE 2 (bases 1 to 4208)
AUTHORS Boger,E.T., Sellers,J.R. and Friedman,T.B.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2001) Laboratory of Molecular Genetics, National Institute on Deafness and Other Communication Disorders, 5 Research Court, Rockville, MD 20850, USA
FEATURES Location/Qualifiers
source 1..4208
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24-q25"
/clone="cDNA2"
/tissue_type="kidney, stomach"
gene <1..>4208
/gene="MYO15B"
/note="myosin XVBP"
/pseudo
BASE COUNT 862 a 1386 c 1275 g 685 t
ORIGIN



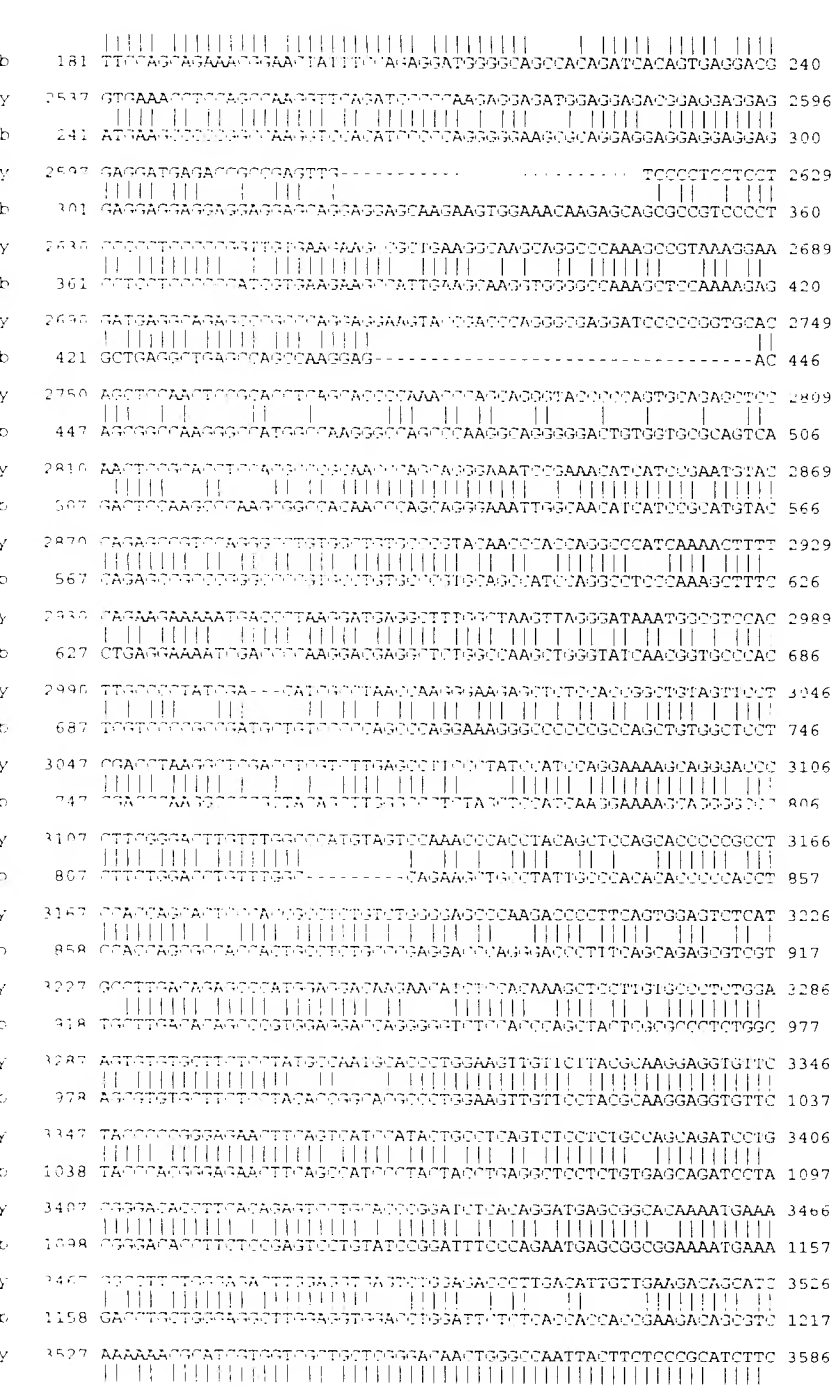


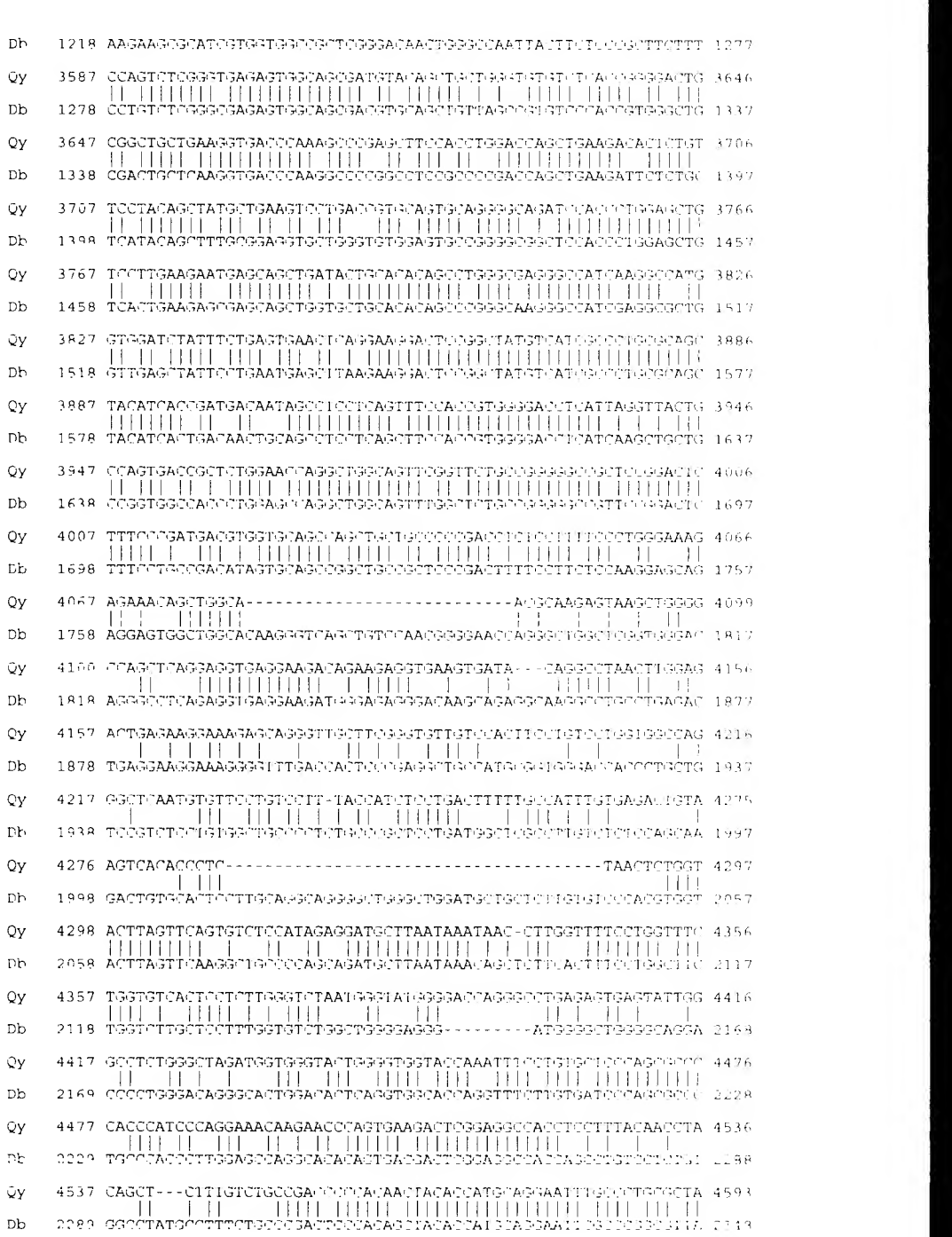


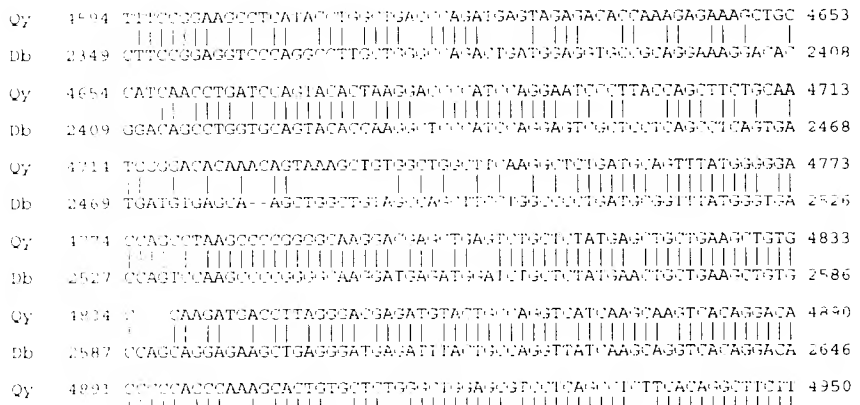


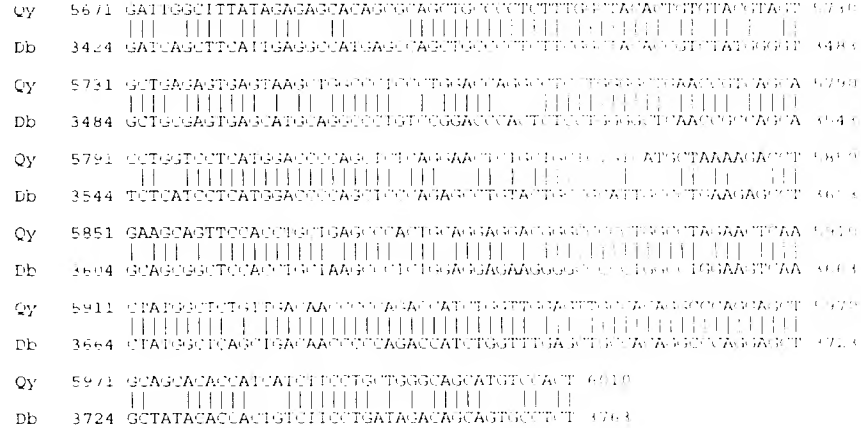


```
RESULT 8
AX253082
LOCUS       AX253082                4174 bp    DNA     linear   PAT 15-SEP-2001
DEFINITION  Sequence 9 from Patent WO0168866.
ACCESSION   AX253082
VERSION     AX253082.1  GI:15986279
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4174)
  AUTHORS   Brooks,A.R., Pemp,C.G. and Rubanyi,G.M.
  TITLE     Estrogen-regulated unconventional myosin related protein,
            compositions and methods of use
  JOURNAL   Patent: WO 0168866-A 9 20-SEP 2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .4174
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Partial DNA sequence for human myosin related
                     protein variant 1 (hMRP1)"
BASE COUNT      901 a   1339 c   1237 g    697 t
ORIGIN
```









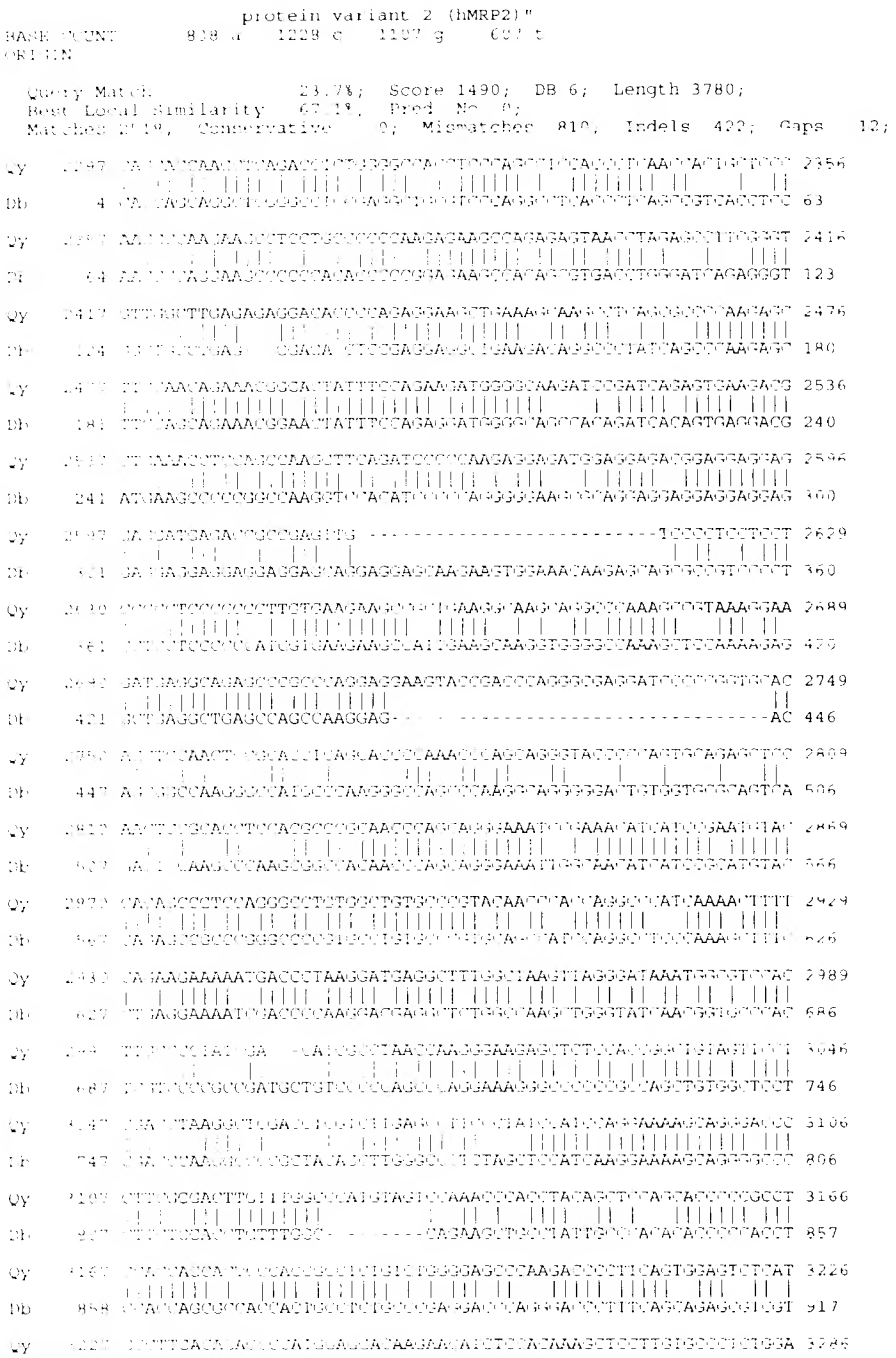
RESULT 9
BC024506
LOCUS BC024506 1547 bp mRNA linear EST 27-AUG-2002
DEFINITION Mus musculus, clone IMAGE:4984596, mRNA, partial cds.
ACCESSION BC024506
VERSION BC024506.1 GI:19353981
KEYWORDS .
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Euarchontoglires, Muridae, Murinae, Mus.
REFERENCE 1 (bases 1 to 1547)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Rm 11A52, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://www.mgc.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305-5329
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdapaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

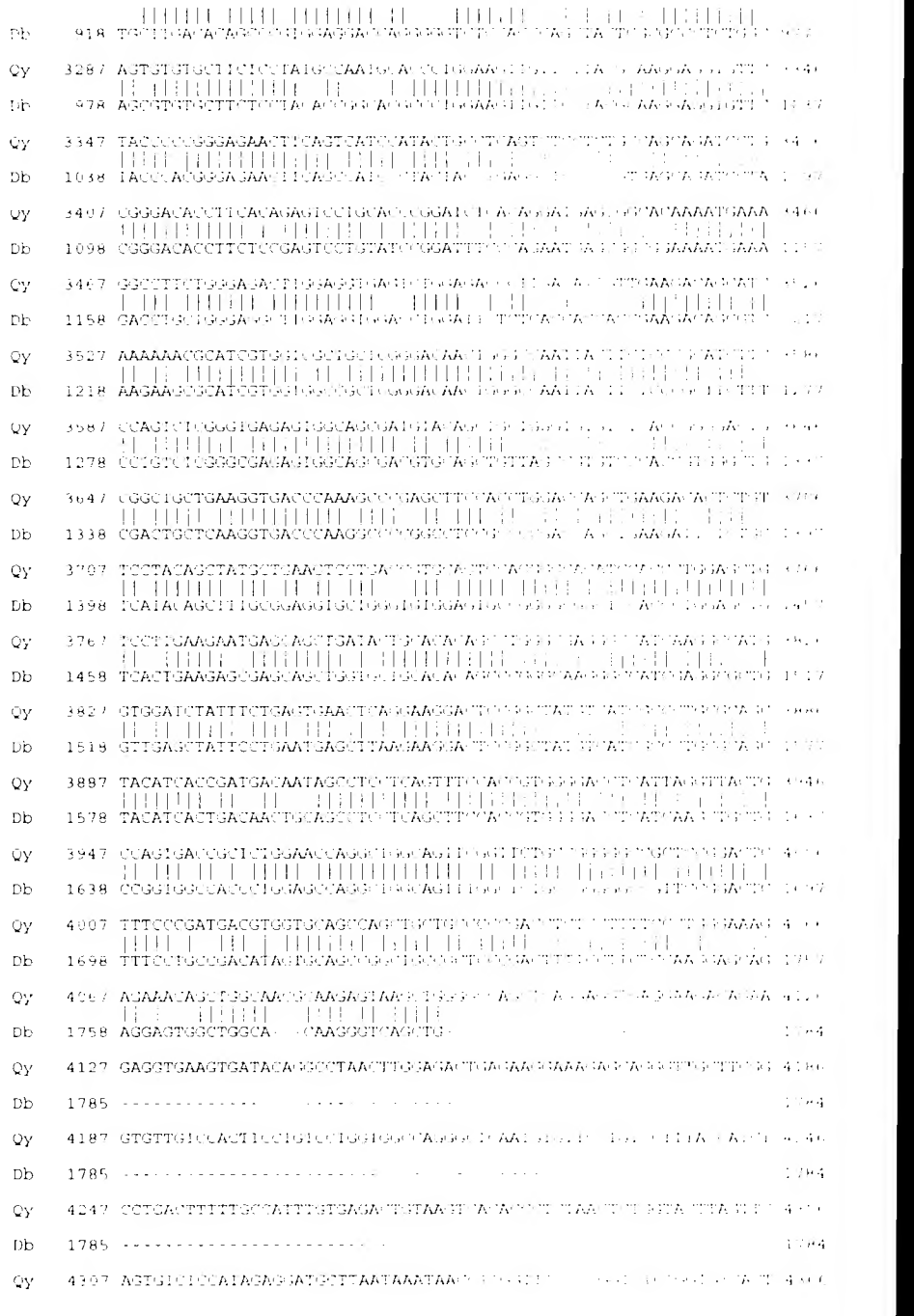
Clone distribution: MGC clone distribution information can be found via the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 59 Row: o Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria. Hexamer frequency DEF analysis.
FEATURES Location/Qualifiers
source 1..1547
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="EVB/N"
/clone="IMAGE:4984596"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6"
CDS <1..1256

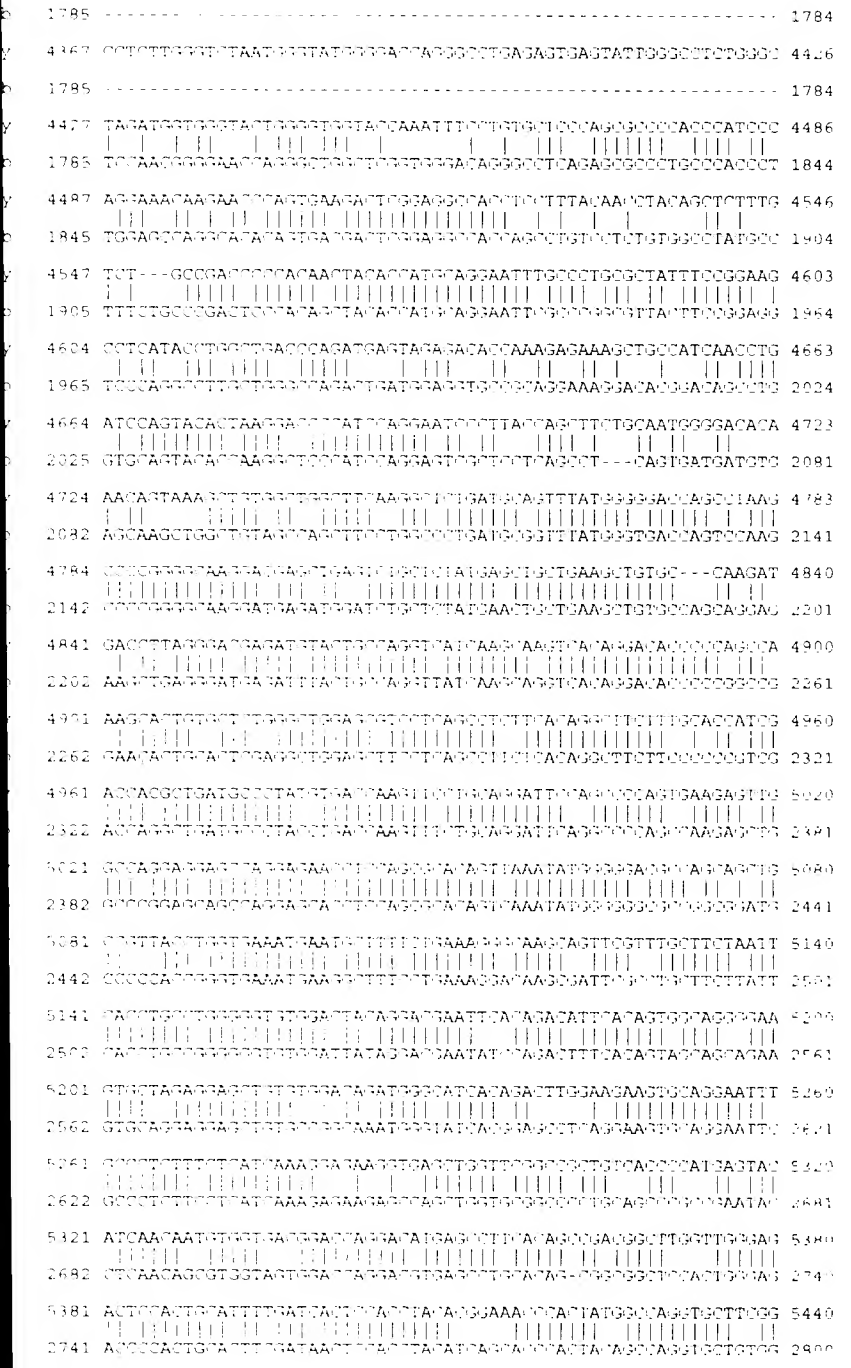
```
                /codon_start=3
                /product="Unknown (protein for IMAGE:4983596)"
                /protein_id="AAH24506.1"
                /db_xref="GI:19353982"
                /translation="PKPPGKDELSLYELLKLCQDDLRDEMYCQVIKQVTGHPQPHC
                ALGWSVLSLFTGFFAPSTTLMPYVIKFLQDSSPSEELARRSQENLQRTVKYGGRQQLP
                LPGEMNAFLKGQAVRLLIHLPGGVDYPTNSQTFTVAGEVLEELCGQMGITDLREVQE
                FALFLIKGEGELVRPLSPHEYINNVVTDQDMSLHSRRLGWETPLHFDHSTYTETHYGQ
                VLPDYLQGKLIVSTQAFALLACLAAPQHFDKTGTSSPPSEQELLSYIPKPLQWQVNTA
                NIKSLVTQEIRQMQGYSKQPAQIGHIESTAQLPLFGYTVYVVLRVSKLALPGPGLLGL
                NRQHLVLMDPSSQPLCCSVMLKDIKQLHLLSPLQFDGPPGLELNYGSVDNFQTIWLEL
                PQAQRLQHIIFLLGSMSTQWPGLL"
BASE COUNT      407 a    429 c    414 g    297 t
ORIGIN

  Query Match              23.8%;  Score 1495.6;  DB 10;  Length 1547;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1504;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   4776 AGCCTAAGCCCCGGGGAAAGAAGGACGAGCTGAGTCTGTCTCTATGAGCTGCTGAAGCTGTGCC 4835
Db      1 AGCCTAAGCCCCGGGGAAAGAAGGACGAGCTGAGTCTGTCTCTATGAGCTGCTGAAGCTGTGCC 60

Qy   4836 AAGATGACCTTAGGGATGAGATGTACTGCCAGGTCATCAAGCAAGTCACAGGACACCCCC 4895
Db     61 AAGATGACCTTAGGGATGAGATGTACTGCCAGGTCATCAAGCAAGTCACAGGACACCCCC 120

Qy   4896 AGCCAAAGCACTGTGCTCTGGGCTGGAGCGTCCTCAGCCTCTTCACAGGCTTCTTTGCAC 4955
Db    121 AGCCAAAGCACTGTGCTCTGGGCTGGAGCGTCCTCAGCCTCTTCACAGGCTTCTTTGCAC 180

Qy   4956 CATCCACCACTCTGATGCCTTATGTGACCAAGTTCCTGCAGGATTCCAGCCCCAGTGAAG 5015
Db    181 CATCCACCACTCTGATGCCTTATGTGACCAAGTTCCTGCAGGATTCCAGCCCCAGTGAAG 240

Qy   5016 AGTTGGCTAGGAGGAGCCAGGAGAACCTGCAGCGCACAGTTAAATATGGGGGACGCCAGC 5075
Db    241 AGTTGGCTAGGAGGAGCCAGGAGAACCTGCAGCGCACAGTTAAATATGGGGGACGCCAGC 300

Qy   5076 AGCTGCCGTTACCTGGTGAAATGAATGCTTTTCTGAAAGGGCAAGCAGTTCGTTTGCTTC 5135
Db    301 AGCTGCCGTTACCTGGTGAAATGAATGCTTTTCTGAAAGGGCAAGCAGTTCGTTTGCTTC 360

Qy   5136 TAATTCACCTTCCTGGGGGTGTGGACTACAGGACGAATTCACAGACATTCACAGTGGCAG 5195
Db    361 TAATTCACCTTCCTGGGGGTGTGGACTACAGGACGAATTCACAGACATTCACAGTGGCAG 420

Qy   5196 TGGAAGTGCTAGAGGAGCTCTGTGGACAGATGGGCATCACAGACTTGGAAGAAGTGCAGG 5255
Db    421 TGGAAGTGCTAGAGGAGCTCTGTGGACAGATGGGCATCACAGACTTGGAAGAAGTGCAGG 480

Qy   5256 AATTTGCCCTCTTTCTCATCAAAGGAGAAGGTGAGCTGGTTCGGCCGCTGTCACCCCATG 5315
Db    481 AATTTGCCCTCTTTCTCATCAAAGGAGAAGGTGAGCTGGTTCGGCCGCTGTCACCCCATG 540

Qy   5316 AGTACATCAACAATGTGGTGACGGACCAGGACATGAGCCTTCACAGCCGACGGCTTGGTT 5375
Db    541 AGTACATCAACAATGTGGTGACGGACCAGGACATGAGCCTTCACAGCCGACGGCTTGGTT 600

Qy   5376 GGGAGACTCCACTGCATTTTGATCACTCCACCTACACGGAAACCCACTATGGCCAGGTGC 5435
Db    601 GGGAGACTCCACTGCATTTTGATCACTCCACCTACACGGAAACCCACTATGGCCAGGTGC 660

Qy   5436 TTCGGGACTACCTGCAAGGCAAGCTGATAGTCAGCACCCAGGCAGAGGCTCTACTTGCCC 5495
Db    661 TTCGGGACTACCTGCAAGGCAAGCTGATAGTCAGCACCCAGGCAGAGGCTCTACTTGCCC 720

Qy   5496 AGCTTGCTGCCTTCCAACACTTCGACAAAACCGGAACTTCTAGTCCTCCATCAGAGCAAG 5555
Db    721 AGCTTGCTGCCTTCCAACACTTCGACAAAACCGGAACTTCTAGTCCTCCATCAGAGCAAG 780

Qy   5556 AGCTGCTGTCTTATATTCCCAAGCCACTGCAATGGCAGGTGAACACAGCCAACATAAAGA 5615
Db    781 AGCTGCTGTCTTATATTCCCAAGCCACTGCAATGGCAGGTGAACACAGCCAACATAAAGA 840

Qy   5616 GCTTGGTGACCCAGGAGCTGAGGCAGATGCAAGGGTACAGCAAGCAGAGAGCACAGATTG 5675
Db    841 GCTTGGTGACCCAGGAGCTGAGGCAGATGCAAGGGTACAGCAAGCAGAGAGCACAGATTG 900

Qy   5676 GCTTTATAGAGAGCACAGCGCAGCTGCCCCTCTTTGGCTACACTGTGTACGTAGTGCTGA 5735
Db    901 GCTTTATAGAGAGCACAGCGCAGCTGCCCCTCTTTGGCTACACTGTGTACGTAGTGCTGA 960

Qy   5736 GAGTGAGTAAGCTGGCCCTCCCTGGACCAGGCCTCCTGGGGCTGAACCGTCAGCACCTGG 5795
Db    961 GAGTGAGTAAGCTGGCCCTCCCTGGACCAGGCCTCCTGGGGCTGAACCGTCAGCACCTGG 1020

Qy   5796 TCCTCATGGACCCCAGCTCTCAGGAACTCTGCTGCTCTGTCATGCTAAAAGACCTGAAGC 5855
Db   1021 TCCTCATGGACCCCAGCTCTCAGGAACTCTGCTGCTCTGTCATGCTAAAAGACCTGAAGC 1080

Qy   5856 AGTTCCACCTGCTGAGCCCACTGCAGGAGGACGGGCCCCCTGGCCTAGAACTCAACTATG 5915
Db   1081 AGTTCCACCTGCTGAGCCCACTGCAGGAGGACGGGCCCCCTGGCCTAGAACTCAACTATG 1140

Qy   5916 GCTCTGTTGACAACCCCCAGACCATCTGGTTGGAGTTGCCACAGGCCCAGGAGCTGCAGC 5975
Db   1141 GCTCTGTTGACAACCCCCAGACCATCTGGTTGGAGTTGCCACAGGCCCAGGAGCTGCAGC 1200

Qy   5976 ACACCATCATCTTCCTGCTGGGCAGCATGTCCACTCAGTGGCCAGGTCTCCTCTGAGGAG 6035
Db   1201 ACACCATCATCTTCCTGCTGGGCAGCATGTCCACTCAGTGGCCAGGTCTCCTCTGAGGAG 1260

Qy   6036 TGGAGATAAGGCAGCGGTCTCTCACTGGGCAGTCTGCCTTAGTCCTGCTCTGAATCCGCT 6095
Db   1261 TGGAGATAAGGCAGCGGTCTCTCACTGGGCAGTCTGCCTTGGTCCTGCTCTGAATCCGCT 1320

Qy   6096 GCACAACCCCCCACCCCACGTGGAGGCCAAAAGGCAAAGTTGTGTCACCTGGGAGAATAG 6155
Db   1321 GCACAACCCCCCACCCCACGTGGAGGCCAAAAGGCAAAGTTGTGTCACCTGGGAGAATAG 1380

Qy   6156 GCAGACACATCCCCTCTGGGGTGGACTGCAACAGGAGTTGGGGCATTTGCTGGCTAGCCC 6215
Db   1381 GCAGACACATCCCCTCTGGGGTGGACTGCAACAGGAGTTGGGGCATTTGCTGGCTAGCCC 1440

Qy   6216 CAGGGAAAATGCCCACCCAGCTCGAAAGCGGCACAAGTAAAACACCCAAGGAAAAAAAAA 6275
Db   1441 CAGGGAAAATGCCCACCCAGCTCGAAAGCGGCACAAGTAAAACACCCAAGAAAAGATTGT 1500

Qy   6276 AAAAAAAAAAAAAAAAAA 6293
Db   1501 GTGTGAAAAAAAAAAAAA 1518


RESULT 10
AX253084
LOCUS       AX253084               3780 bp    DNA     linear   PAT 05-OCT-2001
DEFINITION  Sequence 7 from Patent WO0168866.
ACCESSION   AX253084
VERSION     AX253084.1  GI:15986280
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3780)
  AUTHORS   Brooks,A.R., Deng,G.G. and Rubanyi,G.M.
  TITLE     Estrogen-regulated unconventional myosin-related protein:
            compositions and methods of use
  JOURNAL   Patent: WO 0168866-A 7 20-SEP-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1..3780
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Partial DNA sequence for human myosin related
```

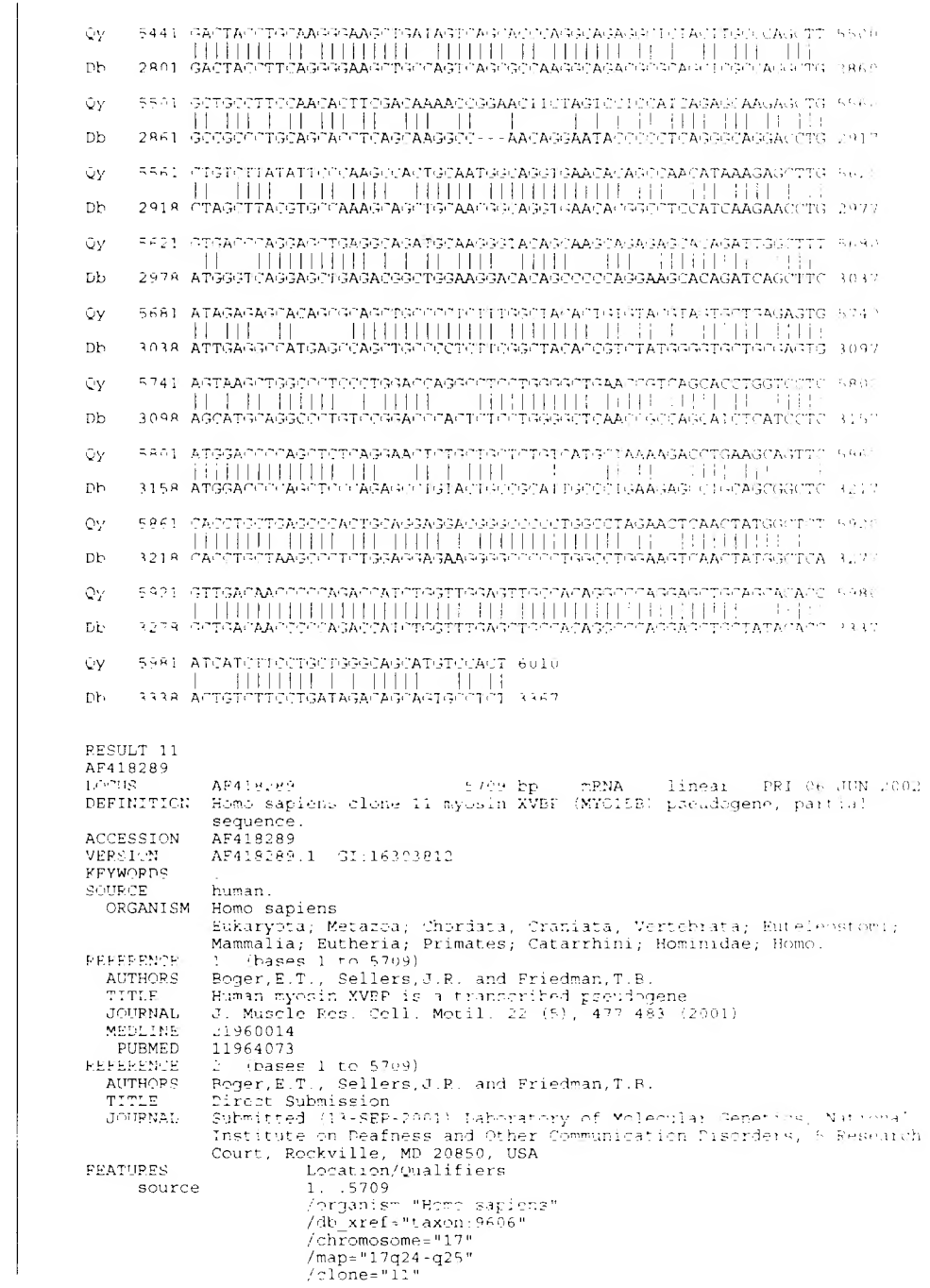
protein variant 2 (hMRP2)"
BASE COUNT 838 a 1228 c 1107 g 607 t
ORIGIN

Query Match 23.7%; Score 1490; DB 6; Length 3780;
Best Local Similarity 67.1%; Pred. No. 0;
Matches 810; Conservative 0; Mismatches 410; Indels 422; Gaps 12;









RESULT 11
AF418289
LOCUS AF418289 5709 bp mRNA linear PRI 06-JUN-2002
DEFINITION Homo sapiens clone 11 myosin XVBP (MYO15B) pseudogene, partial sequence.
ACCESSION AF418289
VERSION AF418289.1 GI:16303812
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5709)
AUTHORS Boger,E.T., Sellers,J.R. and Friedman,T.B.
TITLE Human myosin XVBP is a transcribed pseudogene
JOURNAL J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
MEDLINE 21960014
PUBMED 11964073
REFERENCE 2 (bases 1 to 5709)
AUTHORS Boger,E.T., Sellers,J.R. and Friedman,T.B.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2001) Laboratory of Molecular Genetics, National Institute on Deafness and Other Communication Disorders, 5 Research Court, Rockville, MD 20850, USA
FEATURES Location/Qualifiers
source 1..5709
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24-q25"
/clone="11"

```
                    /tissue_type="stomach"
     gene           <1..>5709
                    /gene="MYO15B"
                    /note="myosin XVBP"
                    /pseudo
BASE COUNT     1041 a   1830 c   2017 g    821 t
ORIGIN

  Query Match             17.8%;  Score 1119.4;  DB 9;  Length 5709;
  Best Local Similarity   75.2%;  Pred. No. 5.8e-250;
  Matches 14[illegible];  Conservative    0;  Mismatches  416;  Indels   71;  Gaps    4;

Qy   685 [illegible]CCAG[illegible]AAGCCATATC[illegible]AGCGGAGGTCAGCTCTGGGACAGCTGAACACCATTCT 744
Db  3752 [illegible]AAAGCAAGCGGTACCTCGGCGGCGGGTAGCTTTGGGACAGCTGAACACCATCCT 3811

Qy   745 [illegible]TAGTGGCCCGGCCCCTGCTCCGGAGA[illegible]ACAGAAGCTACGGTGTGCCCCTGGCCCG-- 802
Db  3812 [illegible]TGGTGGCCCAGCCCCTGCCCCAGAGGCGGCAGAGACTGCAGCTTGGGCGTTGGCAGGG 3871

Qy   [illegible] [illegible]CACACGGGGAGC[illegible]GGGGAAAGTGTCAAATATGGACCTGGGTCGCTTAGA 855
Db  3872 [illegible]ATAG[illegible]AGCGAAAGGGCCTTGGAGA[illegible]AGTGC[illegible]AAGCATGGAGCTGGGGCGCTTGGA 3931

Qy   856 GAT[illegible]TGGC[illegible]CAGCTGGCTACTCTGCTGGAGAGGGCGGAAGGCCACCAGGCCTTGCTGAC 915
Db  3932 [illegible]AT[illegible]CGGCTGAGCTGGCTGTCATGCTGAAGACGGCGGAAAG[illegible]CATCGGGACGCCCTGGC 3991

Qy   916 [illegible]AC[illegible]CA[illegible]ACAGAC[illegible]CCTCC[illegible]ACCTGAGGTCCCG[illegible]AGC[illegible]TGACTCTCCC 975
Db  3992 T[illegible]GAGCATCACCGAGTGCCTGCCGCCTGAGGTTCCTGCCCGGCCCAGCCTGACTCTCCC 4051

Qy   976 T[illegible]AGACATTGACCAGTT[illegible]CAGTTTGTATCCACCAGCTTTCAGAAGCCATT 1035
Db  4052 A[illegible]AGACATTGACCTGTTCCCTTTCTCCAGCTTCGTCGCCATCGGCTTTCAGGAACCGTC 4111

Qy  1036 T[illegible]CCTCGACCAGGGCAGCCACTGGACGAGCCCCTGACGCGGTTAGATGGCGAGAACCC 1095
Db  4112 [illegible]CCAGG[illegible]AGGGCAGCCACTGGCGAAGCCCCTAACCCAGCTGGATGGAGACAACCC 4171

Qy  1096 T[illegible]AGGCTCTGGAGATCAACAGGGTGATGCTGCGGCTCCTGGGGAAGGATCTCTGCA 1155
Db  4172 T[illegible]TGCCCTGGA[illegible]ATCAACAAAGTGATGCTGCGGCTCCTGGGGACGGATCCCTGGA 4231

Qy  1156 [illegible]TGGCAAGAGCAGACCATGGGCACGTTCCTCGTGCAGCAGGCCCAGCGACGGCCGGG 1215
Db  4232 [illegible]AGA[illegible]CA[illegible]ATGG[illegible]GCATA[illegible]TGG[illegible]GCAGGGGCAGTGCTGGCCAGG 4291

Qy  1216 A[illegible]GA[illegible]GATGAGCT[illegible]TTCAGCCAGCTGGTGGCCCAGCTGTGGCGCAACCCAGATGAGCA 1275
Db  4292 [illegible]GGAATGAGCTCTTTAGCCAGCTGGTGGCCCAGCTATGGCAGAACCCAGATGAACA 4351

Qy  1276 A[illegible]ACAATCAGCGTGGCTGGGCCCTAATGGTGATCCTGCTCAGCTCCTTTGCTCCCACACC 1335
Db  4352 [illegible]A[illegible]AGCCAGCGTGGCTGGGCCTCATGGCTGTTTTGCTCAGCGCCTTTCCCCCACTGCC 4411

Qy  1336 T[illegible]TGGAGAAGCCACTGCTCAAATTTGTATCTGACCAGGCTCCCAGTGGCATGGCAGC 1395
Db  4412 T[illegible]CTACAGAAGCCACTGCTCAAGTTCGTGTCTGACCAGGCCCCCAGGGGCATGGCAGC 4471

Qy  1396 [illegible]CAGCACAAGCTGTTAGGTGCCCTGGAGCAGACAC[illegible]CTGGCT[illegible]CATGGCTTC 1455
Db  4472 [illegible]TGCCAGCACAAGCTGCTGGGGGCCTTGGAGCAGTCGCAGCTGGCCTCGGGGGCCAC 4531

Qy  1456 CA[illegible]CCCACCTACCCACACAACTTGAGTGGAAGGCTGGTTTACGTCGGGGCCGCATGGC 1515
Db  4532 T[illegible]GCCCACCCCCCGACCCAGCTGGAGTGGCTGGCCGGATGGCGGCGGGGCCGCATGGC 4591

Qy  1516 [illegible]TGCATGTGTTCACATTCAACGAGGAAAGCTACTCCGGCGGAAGTGGAATCCTGGACCAC 1575
Db  4592 [illegible]TGGATGTGTTCACCTTCAGCGAGGAGTGCTACTCGGCCGAGGTGGAGTCTTGGACCAC 4651

Qy  1576 G[illegible]AGAGCAGTTTGCAGGGTGGATTCTACAGAGC------------------------- 1609
Db  4652 CGGGCAGCAGCTGGCT[illegible]GGTGGATCCTG[illegible]AGAGCAGG[illegible]AGAAAA[illegible]AAATGC[illegible]AT[illegible] [illegible]

Qy  1610 ---------------------------AGAGGC[illegible]GA[illegible] [illegible]
Db  4712 GAGCCCTTGTGGAAGTGGAGGGCCGTGAGAGGCCTGGAGG[illegible] [illegible]

Qy  1643 GTGTCACTGCATTCTGGGGAT[illegible]TTGG[illegible]GTGACTT[illegible]GA[illegible]TTTGTGTTGGA[illegible] [illegible]
Db  4772 GTGTCACTGCACTCCAGGGACGCATGGCAGGACTTGGCTGGTTGCGACTTTGTGCTGGA[illegible] [illegible]

Qy  1703 CTAATAGGCCAGACTGAGGACTTGGGAGA[illegible]CAG[illegible]A[illegible]AA[illegible]TAC[illegible]CATCACT [illegible]
Db  4832 CTGATCAGCCAGACTGAGGACCTGGGGGACCCAG[illegible]CAGCTACCCCATCACT [illegible]

Qy  1763 CCTCTTGGTTTAGCTGAGAGCATCCCTCCAGCCCT[illegible]GTGT[illegible]A[illegible]T[illegible]TT[illegible]TG[illegible] [illegible]
Db  4892 CCTCTTGGCTCAGCCGAGGCCATTCCCCTTGCCCCTGGCATTCA[illegible]TCACTG[illegible] [illegible]

Qy  1823 CCAGGACTCCCTCCAGGTCCAGC[illegible]CCAATAC[illegible]A[illegible]A[illegible]GA[illegible] [illegible]
Db  4952 CCAGGACCCCTCCAGGTCCAGCCCCAACG[illegible]TGC[illegible]CA[illegible]TG[illegible]GAC[illegible]ACACAGG[illegible]AGG[illegible] [illegible]

Qy  1883 ACTAAGCCTCAGAA[illegible]CTGGATGGTT[illegible]GA[illegible]A[illegible]AA[illegible]A[illegible] [illegible]
Db  5012 CAGAGGTCAGGGAGCCTGGACGGCTTCCTGGACCAGA[illegible]AG[illegible]AGTGATATC[illegible]TCC [illegible]

Qy  1943 GGTTTCAGTGATCTGGAACAAGGCTGG[illegible]GA[illegible]A[illegible]A[illegible]AT[illegible]AA[illegible]AG[illegible]T [illegible]
Db  5072 GGCCTCAGCGATCTGGAACAAAGCTGGGCTCTGAGCAG[illegible]ATGAA[illegible]GGAGGGG[illegible] [illegible]

Qy  2003 GTTGGGCCCACCCAGCAGGGCTAC[illegible]CATGGTGTA[illegible]A[illegible]A[illegible]TTAGG[illegible]ACTTA[illegible] [illegible]
Db  5132 ATTGGGCCCACACAGCAGGGCTACCCCATGGTGTAC[illegible]AGGAATGAT[illegible]AGATG[illegible]A [illegible]

Qy  2063 TACCAGCCAGCTATGATAGCC[illegible]A[illegible]GAT[illegible]CCT[illegible]AT[illegible]A[illegible]AT[illegible]A[illegible]A [illegible]
Db  5192 TACCAGCCAGGCATGGTCCCTGCACCCATG[illegible]CATGAT[illegible]AC[illegible]AA[illegible]A[illegible]AG[illegible]T [illegible]

Qy  2123 ACCATGCCAGCCATCATGGTGC[illegible]AC[illegible]CTAG[illegible]CACAG[illegible]AG[illegible]A [illegible]
Db  5252 GCCATGCCAGCCATGGTGGTGCCGCCGCAGCCACCGC      TT[illegible]CAGCCTGGATGCA [illegible]

Qy  2183 ACCCAGCTGGCACTACAGCAC[illegible]AAAAC[illegible]AT[illegible]AA[illegible]A[illegible]A[illegible]ATT[illegible]AG [illegible]
Db  5306 GGGCAGCTGGCCGTCCAGCAGCAGAACTTTATCCAGCAG[illegible]AG[illegible]TAAT[illegible]CAG [illegible]

Qy  2243 CACATGACCACCCAGCCCATGAG[illegible]TGT[illegible]CAG[illegible]A[illegible]A[illegible]AA[illegible]A[illegible]A[illegible]A [illegible]
Db  5366 CAGATGACAGCCCAGGCCATGTCCCAGTCCCTGGAGCAGCAGATGCAGCAG[illegible]GCAGCAG [illegible]

Qy  2303 CAAGCTCAGACCTCTGGGGCCACCTC[illegible]CAGC[illegible]A[illegible]AA[illegible]A[illegible]T[illegible]AA[illegible] [illegible]
Db  5426 CAGGCTCGGGCCTCCGAGGCTGCGTCCCAGGCCTCAC[illegible]TCA[illegible]GTCACCTCCAA[illegible] [illegible]

Qy  2363 AAGAAGCCTCCTGC[illegible]CAAGAGAAGCCAGAGA[illegible]AA[illegible]A[illegible] [illegible]
Db  5486 AGGAAGCCCCCCACACCCCCGGAGAAGCCACAGC[illegible]GA[illegible]CAGGGTGGCTGC [illegible]

Qy  2423 TTGAGAGAGGACACCCCAGAGGAAGCTGAAACAA[illegible]A[illegible]AA[illegible]A[illegible]TT[illegible]AA [illegible]
Db  5546 CTGAG---GGAGACCTCCGAGGAGGCTGAAGACAGG[illegible]TAT[illegible]AG[illegible]AAGAG[illegible]TTCCAG [illegible]

Qy  2483 CACAAACGGGACTATTTCCACAAGATGCC[illegible]AACA[illegible]A[illegible]A[illegible]AACA[illegible]GAAA [illegible]
Db  5603 CAGAAACGGAACTATTTCCAGAGGATGGGCCAGCCACAGAT[illegible]GAGGA[illegible]GATGAA[illegible] [illegible]

Qy  2543 CCTCCAGCCAAGGTTCAGATC[illegible]CAAGA[illegible]GAGA[illegible]GAGGA[illegible]GA [illegible]
Db  5663 CCCCGGCCAAGGTCCACATCCCCCAGGGGGAAGGCAGGAG[illegible]GA 5709
```

RESULT 12

```
K097896
OCUS       AK097896                2372 bp    mRNA    linear   PRI 15-JUL-2002
EFINITION  Homo sapiens cDNA FLJ40577 fis, clone THYMU2007467, weakly similar
           to Mus musculus unconventional myosin-15 mRNA
CCESSION   AK097896
ERSION     AK097896.1  GI:21757796
EYWORDS    oligo capping; fis (full insert sequence).
OURCE      Homo sapiens shymus cDNA to mRNA, clone_lib:THYMU2
           clone:THYMU2007467.
 ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
EFERENCE   1
 AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
           Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
           Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
           Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
           Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
           Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A.,
           Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
           and Isogai,T.
 TITLE     NEDO human cDNA sequencing project
 JOURNAL   Unpublished
EFERENCE   2  (bases 1 to 2372)
 AUTHORS   Isogai,T. and Yamamoto,J.
 TITLE     Direct Submission
 JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
           Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
           (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
MMENT      NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert sequencing:
           Research Association for Biotechnology (RAB); cDNA library
           construction: Helix Research Institute (HRI) (supported by Japan
           Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
           HRI, and Biotechnology Center, National Institute of Technology and
           Evaluation; clone selection for full insert sequencing: HRI and
           RAB; annotation: HRI and RAB.
ATURES              Location/Qualifiers
    source          1..2372
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="THYMU2007467"
                    /tissue_type="shymus"
                    /clone_lib="THYMU2"
                    /note="cloning vector: pME18SFL3"
SE COUNT      486 a    780 c    715 g    391 t
IGIN

Query Match            17.6%;  Score 1110.6;  DB 9;  Length 2372;
Best Local Similarity  72.2%;  Pred. No. 6.5e-248;
Matches 1502;  Conservative    0;  Mismatches  524;  Indels   27;  Gaps    5;

  2031 TGGTGTACCCAGGTATGGTGCAGGCACCTAGCTACCAGCCAGCTATGATACCCGCACCGA 2090
     1 TGGTGTACCCAGGAATGATTCAGATGCCTGCATACCAGCCAGGCATGGTCCCTGCACCCA 60

  2091 TGCCCGTCATGCCAGCCATGGGGCAGTCCCAACCATGCCAGCCATGATGGTGCCACCCC 2150
    61 TGCCCATGATGCCAGCAATGGGCACAGTCCCTGCCATGCCAGCCATGGTGGTGCCGCCGC 120

  2151 AGCCACAGCCTCTGGTGCCCAGTTTGGACTCAAGGCAGCTGGCACTACAGCAGCAAAACT 2210
   121 AGCCACGGC------TTCCAGCCTGGATGCAGGGCAGCTGGCCGTCCAGCAGCAGAACT 174

  2211 TCATCAACCAGCAGGCGATGATTCTGGCGCAGCAGATGACCACCCAGGCCATGAGCCTGT 2270
   175 TTATCCAGCAGCAGGCGCTAATCCTGGCCCAGCAGATGACAGCCCAGGCCATGTCCCTGT 234

  2271 CCCTGGAGCAGCAGAATCAGAGACACCAGCACCAAGCTCAGACCTCTGGGGCCACCTCCC 2330
   235 CCCTGGAGCAGCAGACGCAGCAGCGGCAGCAGCAGGCTCGGGCCTCCGAGGCTGCGTCCC 294

  2331 AGCCTCCACCCTCAACCACTGCTCCCAAGGCCAAGAAGCCTCCTGCCCCCCAAGAGAAGC 2390

Db     295 AGGCTTCACCCTCAGCCTCACCTCCAAGGCCAGGAAGCCCCCTCACACCCCCAGGAACC 354

Qy    2391 CAGAGAGTAACCTAGAGCCTTCGGGTGTTCGCTTGAGAGAGGACACCCCAGAGGAAG--- 2450
Db     355 CACAGCCGTGACCTGGGATCAGAGGGTGGCTGCCTGAG---GGAGACCTCCGAGGAGGCTG 411

Qy    2451 AAAGCAAGCCTCAGCGCCCAAGAGCTTCCAACAGAAACGGGACCATTTCCAGAAGATGG 2510
Db     412 AAGACAGGCCCTATCAGCCCAAGAGCTTCCAGCAGAAACGGAACTATTTCCAGAGGATGG 471

Qy    2511 GGCAAGATCCGATCAGAGTGAAGACGGTGAAACCTCCAGCCAAGGTTCAGATCCCCCAAG 2570
Db     472 GGCAGGCACAGATCACAGTGAGGACGATGAAGCCCCCGGCCAAGGTCCACATCCCCCAGG 531

Qy    2571 AGGAGATGGAGGAGACGGAGGAGGAGGAGGATGAGACGGCGAGTTGTCCCCTCCTCCTT 2630
Db     532 GGGAAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCAGGAGGAGCAAGAAG 591

Qy    2631 CCCCTCCCCCGGTTGTGAAGAAGCCGCTGAAGGCAAGCAGCCCCAAAGCCCTAAAGGAAG 2690
Db     592 TGGAAACAAGAGCAGTGCCGTCCCCTC----CTCCTCCCCCCATCGTGAAGAAGCCAT 645

Qy    2691 ATGAGGCAGAGCCCGCCCAGGAGGAAGTACCGACCCAGGGCGAGGATCCCCCGGTGCACA 2750
Db     646 TGAAGCAAGGTGGGGCCAAAGCTCCAAAAGAGGCTGAGGCTGAGCCACCCAAGGAGACAG 705

Qy    2751 GCTCCAACTCCGCACCTCAGCACCCCAAACCCAGCAGGGTACCCCCAGTGCAGAGCTCCA 2810
Db     706 CGGCCAAGGGCCATGGCCAAGGGCCAGCCCAAGGGCAGGGGACTGTGGTGCCCAGCTCAG 765

Qy    2811 ACTCCGCACCTCACGCCCGCAACCCAGCAGGAAATCCGAAACATCATCCGAATGTACC 2870
Db     766 ACTCCAAGCCTAAGCCGCCACAACCCAGCAGGAAATCGGAACATCATCCGCATGTACC 825

Qy    2871 AGAGCCGTCCAGGGCCTGTGGCTGTGCCCGTACAACCCACCAGGCCATCAAAACTTTTC 2930
Db     826 AGAGCCCCCCGGGCCCTGTGCCTGTGCCCGTGCAGCCACCCAGGCCTCCAAACCTTTC 885

Qy    2931 AGAAGAAAAATGACCCTAAGGATGAGGCTTTGGCTAAGTTAGGGATAAATGAGGTCCACT 2990
Db     886 TGAGGAAAATCGACCCCAAGGACGAGGCTCTGGCCAAGCTGGGTATCAACGGTGCCCACT 945

Qy    2991 TGCCCCTATCGA---CATCGGCTAACCAAGGGAAGAGCTCTCCACCGGCTGTAGTTCCTC 3047
Db     946 CGTCCCCGGCGGATGCTGTCCCCCAGCCCAGGAAAGGGCCCCCCCGCCCAGCTGTGGCTCCTC 1005

Qy    3048 GACCTAAGGCTGGACCTCGTCTTGAGCCTTCCCTATCCATCCAGGAAAAGCAGGGACCCC 3107
Db    1006 GACCCAAGGCCCCGCTACAGCTTGGGCCCTCTAGCTCCATCAAGGAAAAGCAGGGGCCCC 1065

Qy    3108 TTCGGGACTTGTTTGGCCCATGTAGTCCAAACCCACCTACAGCTCCAGCACCCCCGCCTC 3167
Db    1066 TTCTGGACCTGTTTGGCCAG----------AAGCTGCCTATTGCCCACACACCCCCACCTC 1116

Qy    3168 CACCAGCACTCCCACCGGCCTCTGTCTGGGGAGCCCAAGACCCCTTCAGTGGAGTCTCATG 3227
Db    1117 CACCAGCGCCACCACTGCCTCTGCCCGAGGACCCAGGACCCTTTCAGCAGAGCGTCGTT 1176

Qy    3228 CCTTGACAGAGCCCATGGAGGACAAGAACATCTCCACAAAGCTCCTTGTGCCCTCTGGAA 3287
Db    1177 GCTTGACACAGCCCGTGGAGGACCAGGGGTCTCCACCCAGCTACTCGGCCCTCTGGCA 1236

Qy    3288 GTGTGTGCTTCTCCTATGCCAATGCACCCTGGAAGTTGTTCTTACGCAAGGAGGTGTTCT 3347
Db    1237 GCGTGTGCTTCTCCTACACCGGCACACCCTGGAAGTTGTTCCTACGCAAGGAGGTGTTCT 1296

Qy    3348 ACCCCCGGGAGAACTTCAGTCATCCATACTGCCTCAGTCTCCTCTGCCAGCAGATCCTCC 3407
Db    1297 ACCCACGGGAGAACTTCAGCCATCCCTACTACCTGAGGCTCCTCTGTGAGCAGATCCTAC 1356

Qy    3408 GGGACACCTTCACAGAGTCCTGCACCCGGATCTCACAGGATGAGCGGCACAAAATGAAAG 3467
```

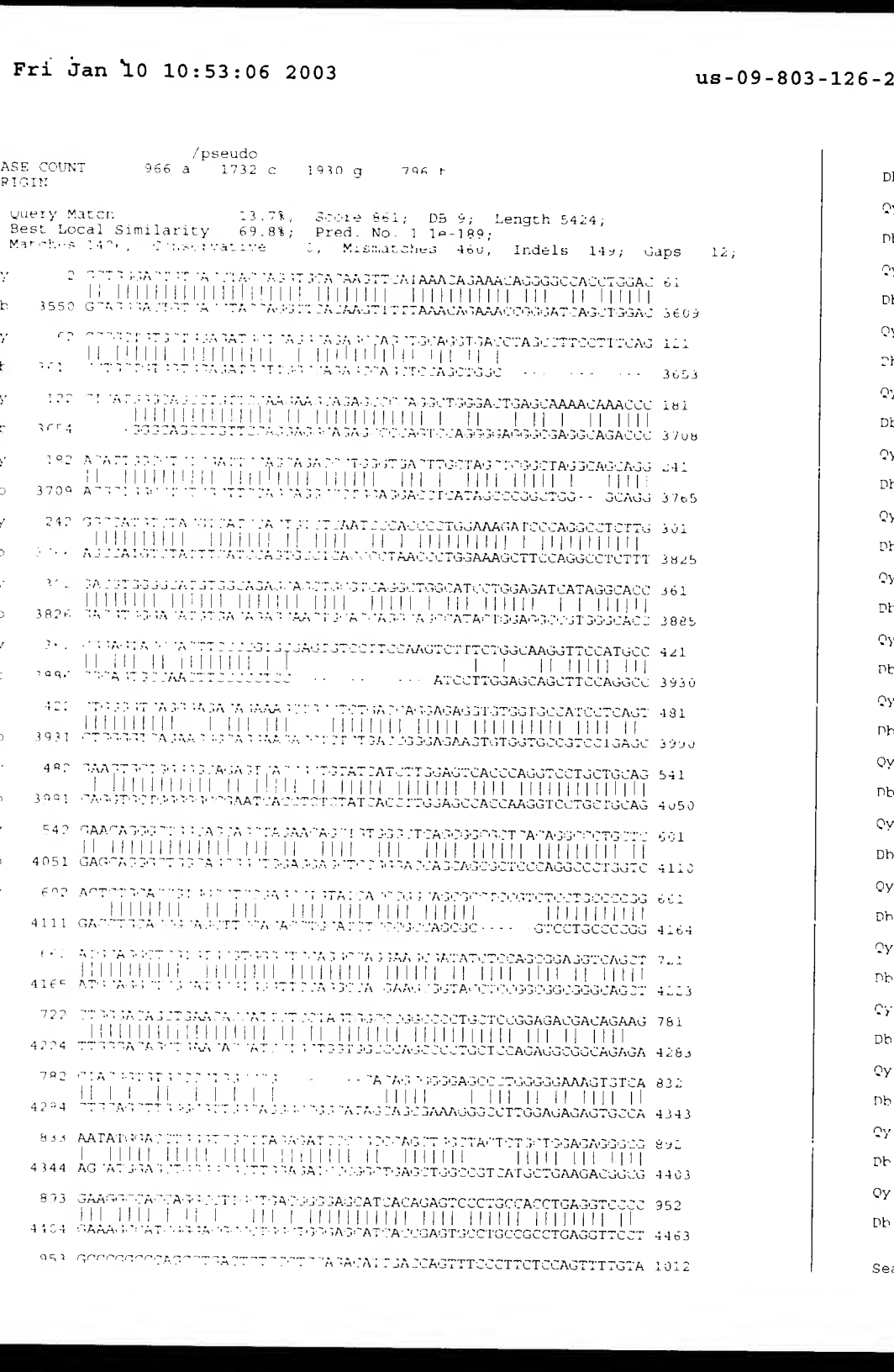
```
Db   1357 GGGACACTTCTC[illegible]CAG[illegible]GTATC[illegible]GGATTTCCCAGAATGAGCGGCGGAAAATGAAAG 1416

Qy   3468 [illegible]GGACACTTGGAGGTGAGTCTGGA[illegible]ACCCTTGACATTGTTGAAGACAG[illegible]ATTA 3527
Db   1417 AC[illegible]TGGGAGGCTTGGAGGTGGACCTGGATTCTCTCACCACCACGAAGACAGGGTCA 1476

Qy   3528 AAAAAG[illegible]CATGGTGGTC[illegible]TGCTG[illegible]GGACAACTGGGCCAATTACTTCT[illegible]AT[illegible]TTC[illegible] 3587
Db   1477 AAAA[illegible]CCATCGTGGTG[illegible]GCTGGGATAACTGGGCCAATTACTT[illegible] 1536

Qy   3588 CA[illegible]GGCTGAGAGTG[illegible]AGGCATGTACAGCTGCTGGGTGTGTCTCACGGGGGACTGC 3647
Db   1537 [illegible]TCGGGCGGGAGTGGCAGCGACGTGCAGCTGTTAGCCGTGTCCCACCGTGGGCTGC 1596

Qy   3648 [illegible]TGAAG[illegible]GA[illegible]AAAG[illegible]GA[illegible]TGGA[illegible]CAGCTGAAGACACTCTGTT 3707
Db   1597 GA[illegible]TCAAGGTCACCCAAGG[illegible]GACCAGCTGAAGATTCTCTGCT 1656

Qy   3708 [illegible]TACAGCTATGCTGAAGTCCTGACCGTGCAGTGCAGGGGCAGATCCACCCTGGAGCTGT 3767
Db   1657 CATA[illegible]AGCTTCGGCGGAGGTGCTGGCTGTGGAATG[illegible]GGGGCGGCTCCACCCTGGAGCTGT 1716

Qy   3768 [illegible]AAGAA[illegible]CAGCAG[illegible]GATA[illegible]TACATAGTTTGG[illegible]AGG[illegible]ATTAAG[illegible]ATGG 3827
Db   1717 CA[illegible]TGAAGAG[illegible]GAGCAG[illegible]ACATAG[illegible]GGGCAAGGGCCATCGAGGGGCTGG 1776

Qy   3828 T[illegible]ATCTATTTCTGAGTGAACTCAGGAAGCACTCCGGCTATGTCATGGCCCTGGGCAGCT 3887
Db   1777 TTGA[illegible]TATTC[illegible]TGAACGAGCTTAAGAAGGACTCCGGCTATGTCATGGCCCTGGGCAGCT 1836

Qy   3888 ACAT[illegible]ATCGATGACAATA[illegible]TTTTAGTTTTCACC[illegible]ACCTCATTAGGTTACTGC 3947
Db   1837 AC[illegible]TACTCACAAT[illegible]AC[illegible]GGACT[illegible]ATTAAGCTGCTGC 1896

Qy   3948 [illegible]A[illegible]GACCG[illegible]TGGAA[illegible]GGATTTT 4007
Db   1897 [illegible]GGAG[illegible]AGGCTGGTAGTTTGGCTTTG[illegible]GGG[illegible]GTT[illegible]AC[illegible] 1956

Qy   4008 TT[illegible]ATGACGTGGTGCAGCCAGCTGCTGCCCCGACCTCTCCTTTCCCTGGGAAAGA 4067
Db   1957 TTCCTGCCGACATAGTGCAGCCGGCTGCCGCTCCCGACTTTTCCTTCTCCAAGGAGCAGA 2016

Qy   4068 GAAACAGCTGGCA 4080
Db   2017 GGAGTGGCTGGCA 2029


RESULT 1[illegible]
BC0278[illegible]
LOCUS       BC027875               2347 bp    mRNA    linear   PRI 01-MAY-2002
DEFINITION  Homo sapiens, hypothetical protein FLJ22686, clone MGC:34148
            IMAGE:5205535, mRNA, complete cds.
ACCESSION   BC027875
VERSION     BC027875.1  GI:20379900
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2347)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L.,
            Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
            Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C.,
            Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 49 Row: g Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: [illegible]
FEATURES             Location/Qualifiers
     source          1..2347
                     /organism="Homo sapiens"
                     /db_xref="LocusID:80022"
                     /db_xref="[illegible]"
                     /clone="MGC:34148 IMAGE:5205535"
                     /tissue_type="Lung, Spleen, fetal, pooled"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6"
     CDS             118..1881
                     /codon_start=1
                     /product="hypothetical protein FLJ22686"
                     /protein_id="AAH27875.1"
                     /db_xref="GI:20379901"
                     /translation="MQEFARRYFRRS[illegible]AAGK[illegible]
                     [illegible]AVA[illegible]ALM[illegible]
                     QVTGHPRPEHCTRGWSFLSLLTGFFPPSTELMPYLTKFLQDSGPSQELARSQ[illegible]
                     [illegible]MEAL[illegible]QAIR[illegible]VAAEVQEE[illegible]
                     MGITEPQEVQEEALELIKEFS[illegible]
                     [illegible]
                     KQLQRQVNIA[illegible]
                     ALSGPILLGLNRQHLILMDLGSQSLT[illegible]
                     DNPQTIWFELPQAQELLYTTVFLIDSSAS[illegible]TEW[illegible]N"
BASE COUNT      574 a    719 c    654 g    400 t
ORIGIN

  Query Match             14.9%;  Score 936.8;  DB [illegible];  Length 2347;
  Best Local Similarity   74.0%;  Pred. No. [illegible];
  Matches 1275;  Conservative    0;  Mismatches  4[illegible];  Indels  [illegible];  Gaps  [illegible];

Qy   4294 TGGTACTTAGTTCAGTGTCTCCATAGAGGATG[illegible]TTAATAAATAA[illegible] 4[illegible]
Db    146 TGGTACTTAGTTCAAGG[illegible]TGCC[illegible]AGTAGAT[illegible]TTAATAAA[illegible] 2[illegible]

Qy   4353 TTTCTGGTGT[illegible]ACTCC[illegible]AA[illegible]TGA[illegible]ACTGACTA 4412
Db    206 CTTCTGGTCTTGCTCCTTTGGTGTCTGGCTGGGGA[illegible]GA [illegible] 2[illegible]

Qy   4413 TT[illegible]AAA[illegible] 447[illegible]
Db    257 AGGACCCCTGGGACAGGG[illegible]ACTGGACACT[illegible] 3[illegible]

Qy   4473 GCCCCACCCATCCCAGGAAACAAGAAC[illegible]AGTGAAGA[illegible]TTTACAA 45[illegible]
Db    317 GCCCTGC[illegible] 3[illegible]

Qy   4533 CCTACAGCT--CTTTGTCTGCCGACCCCCACAACTACACCATG[illegible]AGGAATTT[illegible] 45[illegible]
Db    377 CTGTGGCCTATGCCTTTCTGCCCGACTC[illegible]AATT[illegible] 43[illegible]

Qy   4590 G[illegible]TATTTCTGGAAGCCTCATAC[illegible] 46[illegible]
Db    437 GTTACTTCCGGAGCT[illegible]CAGGCCTTCCTG[illegible] 49[illegible]
```

```
Qy  4650 [illegible] 4709
Db   497 [illegible] 554
Qy  4710 [illegible] 4769
Db   555 [illegible] 613
Qy  4770 [illegible] 4829
Db   614 [illegible] 673
Qy  4830 [illegible] 4886
Db   674 [illegible] 733
Qy  4887 [illegible] 4946
Db   734 [illegible] 793
Qy  4947 [illegible] 5006
Db   794 [illegible] 853
Qy  5007 [illegible] 5066
Db   854 [illegible] 913
Qy  5067 [illegible] 5126
Db   914 [illegible] 973
Qy  5127 [illegible] 5186
Db   974 [illegible] 1033
Qy  5187 [illegible] 5246
Db  1034 [illegible] 1093
Qy  5247 [illegible] 5306
Db  1094 [illegible] 1153
Qy  5307 [illegible] 5366
Db  1154 [illegible] 1213
Qy  5367 [illegible] 5426
Db  1214 [illegible] 1273
Qy  5427 [illegible] 5486
Db  1274 [illegible] 1333
Qy  5487 [illegible] 5546
Db  1334 [illegible] 1390
Qy  5547 [illegible] 5606
Db  1391 [illegible] 1450
Qy  5607 [illegible] 5666
Db  1451 [illegible] 1510
Qy  5667 [illegible] 5726
Db  1511 [illegible] 1570
Qy  5727 [illegible] 5786
Db  1571 [illegible] 1630
Qy  5787 [illegible] 5846
Db  1631 [illegible] 1690
Qy  5847 [illegible] 5906
Db  1691 [illegible] 1750
Qy  5907 [illegible] 5966
Db  1751 [illegible] 1810
Qy  5967 [illegible] 6010
Db  1811 AGCTGCTATACACCACTGTCTTCCTGATAGACAGCAGTGCCTCT 1854
```

RESULT 14
AK026339

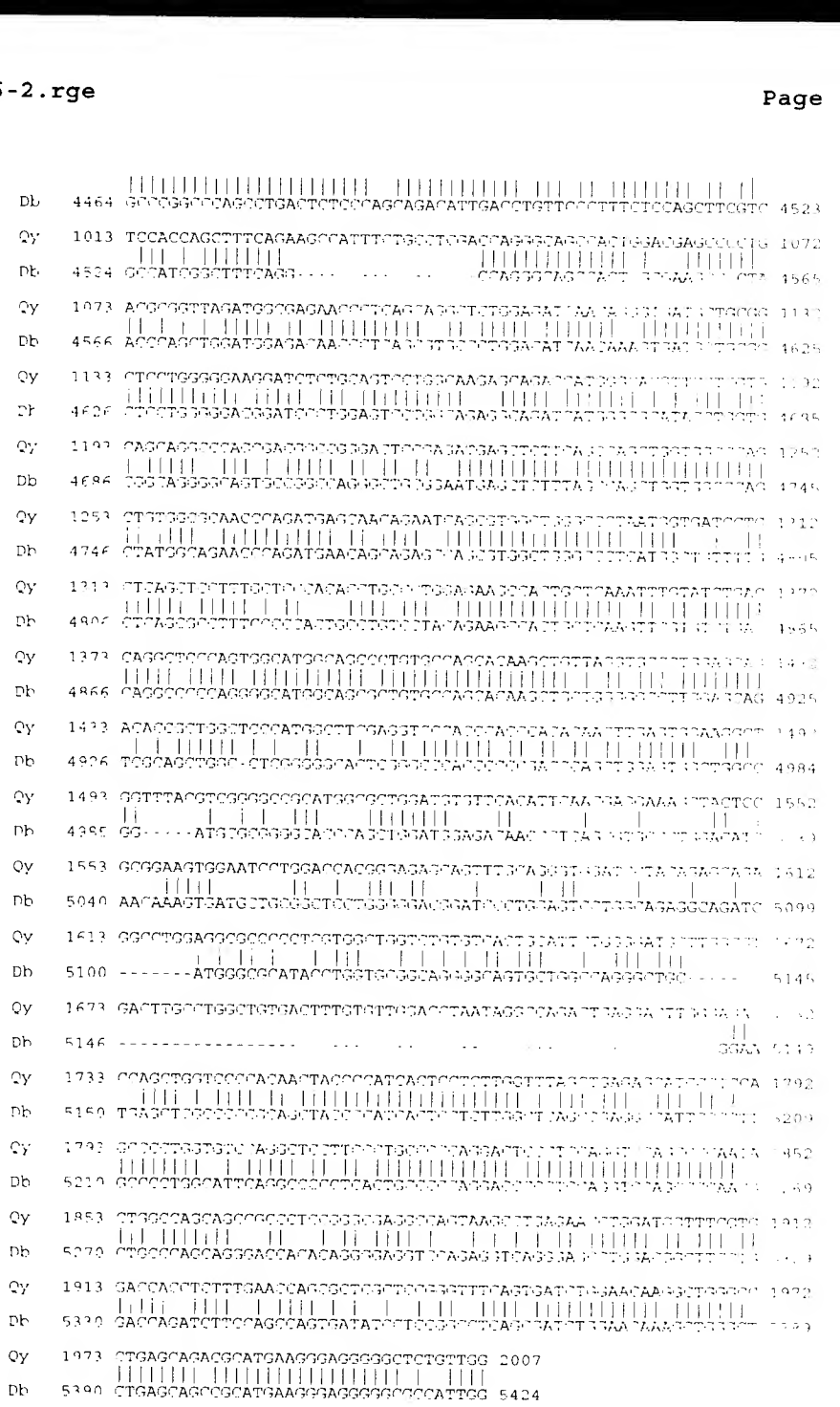
```
LOCUS       AK026339                1891 bp    mRNA    linear   PRI [illegible]-2000
DEFINITION  Homo sapiens cDNA: FLJ22696 fis, clone HSI10987.
ACCESSION   AK026339
VERSION     AK026339.1  GI:10439173
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens human small intestine cDNA to mRNA, clone_lib:HSI
            clone:HSI10987.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1891)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Department of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1..1891
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI10987"
                     /tissue_type="human small intestine"
                     /clone_lib="HSI"
                     /note="cloning vector pME18SFL3"
     CDS             29..1492
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15453.1"
                     /db_xref="GI:10439174"
                     /translation="MQEFARRYFPRSQALLGQTDGGAAGKDTDSLVQYTKAPIQESLL
                     SLSDDVSKLAVASFLALMKFM[illegible]QSKPRGKDEM[illegible]LYELLKL[illegible]EKLR[illegible]EIYCQVIK
                     QVTGHPRPEHCTRGWSFLSLLTGFFPPSTRLMPYLTKFLQDSGPSQELARSSQEHLQR
                     TVKYGGRRRMPPPGEMKAFLKGQAIRLLLIHLPGGV[illegible]YRTNIQTFTVAAEVQEEL[illegible]KQ
                     MGITEPQEVQEFALFLIKEKSQLVRPLQPAECLNSVVVDQDVSLHSRRLHWETPLHFD
                     NSTYISTHYSQVLWDYLQGKLPVSAKADAQLARLAALQHLSKANRNTPSGQDLLAYVP
```

```
                    KQLQRQVNTASIKNLMCQ[illegible]RFLFGHSPQ[illegible]AQISFIFAMSQ[illegible]PLFGYTVYGVLRVSMQ
                    ALGGPIL[illegible]LNR[illegible]HLILM[illegible]CLY[illegible]IALFS[illegible]HL[illegible]LEEF[illegible]EVYG[illegible]A
                    DNPQTIWFELFQAQELLYTTVFLIDSSASCTEWPSIN"
BASE COUNT       454 a    595 c    536 g    306 t
ORIGIN

  Query Match              13.9%;  Score 872.6;  DB 9;  Length 1891;
  Best Local Similarity    76.2%;  Pred. No. 2.1e-192;
  Matches 1115;  Conservative    0;  Mismatches  339;  Indels    9;  Gaps    3;

Qy   4551 [illegible]TACCCCCACAACTACACCATGCAGGAATT[illegible]GGCTATTT[illegible]GGAAGCCTCATA 4610
Db      9 [illegible]GA[illegible]CCCACAGCTACACCATGCAGGAATTCGCCCGG[illegible]TTACTT[illegible]GGAGGT[illegible]CAGG 68

Qy   4611 [illegible]TGACCCAGATGAG[illegible]AGAGACACCAAAGAGAAAGCTGCCATCAACCTGATCCAGT 4670
Db     69 CCTTGCTGGGCCAGACTGATGGAGGTG[illegible]AGGAAAGGACACGGACAGCCTGGTGCAGT 128

Qy   4671 A[illegible]TAAGGACC[illegible]CATCCA[illegible]GAATCCCTTACCAGCTTCTGCAATGGGGACACAAACAGTA 4730
Db    129 A[illegible]AAGGCTCCCATCCAGGAGTGGC[illegible]CAGCCT---CAGTGATGATGTGAGCAAGC 185

Qy   4731 AA[illegible]TCTGCC[illegible]CCTTCAAGGCTCTGA[illegible]CAGTTTAT[illegible]GGGACCAGCCTAAGCCCCGGG 4790
Db    186 T[illegible]TGTAGCCAGCTTCCTGCCCCTCA[illegible]GGTTTATGGGTGACCAGTCCAAGCCCCGGG 245

Qy   4791 [illegible]AAGGACGAGCTGACTCTG[illegible]TATGAGCTGCTGAAGCTGTGC---CAAGATGACCTTA 4847
Db    246 [illegible]AAGGATGAGATGGATCTGCTCTATGAACTGCTGAAGCTGTGCCAGCAGGAGAAGCTGA 305

Qy   4848 [illegible]A[illegible]GAGATGTACTGCCAGGTCATCAAGCAAGTCACAGGACACCCCCAGCCAAAGCACT 4907
Db    306 GGGATGAGATTTACTGCCAGGTTATCAAGCAGGTCACGGGACACCCC[illegible]GGCCGGAACACT 365

Qy   4908 [illegible]TCTGGC[illegible]GGAGCCTCCTCAGCCTCTTCACAGGCTTCTTTGCACCATCGACCAGGC 4967
Db    366 GCA[illegible]GAGG[illegible]TGGAGCTTC[illegible]CAGCC[illegible]CACAGGC[illegible]TTCCCCCGTCGACCAGGC 425

Qy   4968 TGATGCCCTATGTGACCAAGTTCCTGCAGGATTCCAGC[illegible]AGTGAAGAGTTGGCCAGGA 5027
Db    426 T[illegible]CCCTACCTGACCAAGTTTCTGCAGGATTCAGGCCCCAGC[illegible]AAGAG[illegible]TGGC[illegible]GGA 485

Qy   5028 [illegible]CCAGGAGAACCTCCAGCGCACAGTTAAATATGGGGGACGCCAGCAGCTGCGGTTAC 5087
Db    486 [illegible]CAGGAGCACCTCCAGCGCACAGTCAAATATGGGG[illegible]GGATG[illegible]AC 545

Qy   5088 [illegible]GTGAAATGAATGCTTTCTGAAAGGGCAAGCAGTTCGTTTGCTTCTAATTCACCTGC 5147
Db    546 CGGGTGAAATGAAGGCTTTCCTGAAAGGACAAGCGATT[illegible]GCTTCTTATTCACCTGC 605

Qy   5148 [illegible]GGTCTGCACTACACGACGAATTTACAGACATTCACAGTGGCAGGGGAAGTGCTAG 5207
Db    606 [illegible]GGCTGTGGATATAGGACGAATATCCAGACTT[illegible]ACAGTAGCAGCAGAAGTGCAGC 665

Qy   5208 A[illegible]AGCTGTGTGGACAGATGGGCATCACAGACTTGGAAGAAGTGCAGGAATTTGC[illegible] 5267
Db    666 A[illegible]AGCTGTGCCGGCAAATGGGTATCACGGAGCCCAGGAAGTGCAGGAATTCGCCCTCT 725

Qy   5268 TT[illegible]CATCAAAGCAGAAGGTGAGCTGGTTTGGCCGCTGTCACCCCATGAGTACATCAACA 5327
Db    726 T[illegible]TCATCAAAGAGAAGAGCCAGCTGGTGCGG[illegible]GAATGCTTTAATA 785

Qy   5328 [illegible]GAGGAGACTCCAC 5387
Db    786 GCGTGGTAGTGGACCAGGACGTGAGCCTGCACAGCCGGCGGCTCCACTGGAGAC[illegible]AC 845

Qy   5388 T[illegible]ATTTCATCACTCCACCTACAGGAAACCCACTATGGCCAGGTGCTTCGGGACTACC 5447
Db    846 T[illegible]ACTTCGATAACTCCACCTACATCAGCACCCACTACAGCCAGGTGCTGTGGGACTA[illegible]

Qy   5448 T[illegible]AAGGCAAGCTGATAGCTCAGCACCCAGGCAGAGGCTCTACTTGCCCAGCTTCCTGCCT 5507
Db    906 TT[illegible]AGGGGAAGCTGCCAGTCAGCGG[illegible]AAGGCAGAC[illegible]AGCTGGCCAGGCTGGCCGCCC 965

Qy   5508 [illegible]AA[illegible]A[illegible]AAAA[illegible]GAA[illegible]
Db    966 TGCAGCACCTCAGCAAGGCC---AA[illegible]AGGAATA[illegible]

Qy   5568 ATATTCCCAAGCCACTGCAATGGCAGGTGAACACAGCCAACATAAAGAGCTTGGTGACC[illegible]
Db   1023 ACGT[illegible]CCAAAGCAGCTGCAACTGCAGGTGAACACAG[illegible]

Qy   5628 AGGAGCTGAGGCAGATGCAAGGGTACAGCAAGCAGAGA[illegible]
Db   1083 AGGAGCTGAGACGGCTGGAAGGACACAGCCC[illegible]CAGGAAGCACAGATCAGCTTCATCGAGC 114[illegible]

Qy   5688 GCACAGCCGCAGCTGCCCCTCTT[illegible]TACACTGGTGTA[illegible]
Db   1143 CCATGAGCCAGCTGCCCCTCTT[illegible]TACAC[illegible]TA[illegible]

Qy   5748 TGGCCCTCCCTGGACCAGGCCTCCTGGGGCTGAA[illegible]
Db   1203 AGGCCCTGTCCGGACCCACTCTCCTGGGGCTCAA[illegible]

Qy   5808 CCAGCTCTCAGGAACTCTGCTGCTCTGTCATGCTAAAAGACCTGAAGCAGTTCCACCTGC[illegible]
Db   1263 CCAGCTCCCAGAGCCTGTACTGCC[illegible]ATTGCC[illegible]GAAGA[illegible]

Qy   5868 TGAGCCCACTGCAGGAGGACGGG[illegible]TAGAA[illegible]AA[illegible]TTGA[illegible]A [illegible]
Db   1323 TAAGCCCTCTGGAGGAGAAGGGG[illegible]TGGCCTGGAAGTCAA[illegible]TATGGCTCAGCTGACA 1382

Qy   5928 ACCCCCAGACCATCTGGTTGGAGTTGCCACAGGC[illegible]AGGA[illegible]
Db   1383 ACCCCCAGACCATCTGGTTTGAGCTGCCACAGG[illegible]AGGA[illegible]TATACA[illegible] 144[illegible]

Qy   5988 TCCTGCTGGGCAGCATGTCCACT 6010
Db   1443 TCCTGATAGACAGCAGTGCCTCT 1465


RESULT 15
AF418288
LOCUS       AF418288                5424 bp    mRNA    linear   PRI [illegible]
DEFINITION  Homo sapiens clone 9 myosin XVBP (MYO15B) pseudogene, partial
            sequence.
ACCESSION   AF418288
VERSION     AF418288.1  GI:16303811
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5424)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Human myosin XVBP is a transcribed pseudogene
  JOURNAL   J. Muscle Res. Cell. Motil. 22 (5), 477-483 (2001)
  MEDLINE   21960014
   PUBMED   11964073
REFERENCE   2  (bases 1 to 5424)
  AUTHORS   Boger,E.T., Sellers,J.R. and Friedman,T.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (13 SEP 2001) Laboratory of Molecular Genetics, National
            Institute on Deafness and Other Communication Disorders, [illegible]
            Court, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..5424
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q24-q25"
                     /clone="9"
                     /tissue_type="kidney"
     gene            <1..>5424
                     /gene="MYO15B"
                     /note="myosin XVBP"
```



/pseudo
ASE COUNT 966 a 1732 c 1930 g 796 t
RIGIN

Query Match 13.7%; Score 861; DB 9; Length 5424;
Best Local Similarity 69.8%; Pred. No. 1.1e-189;
Matches 1421; Conservative 0; Mismatches 460; Indels 149; Gaps 12;

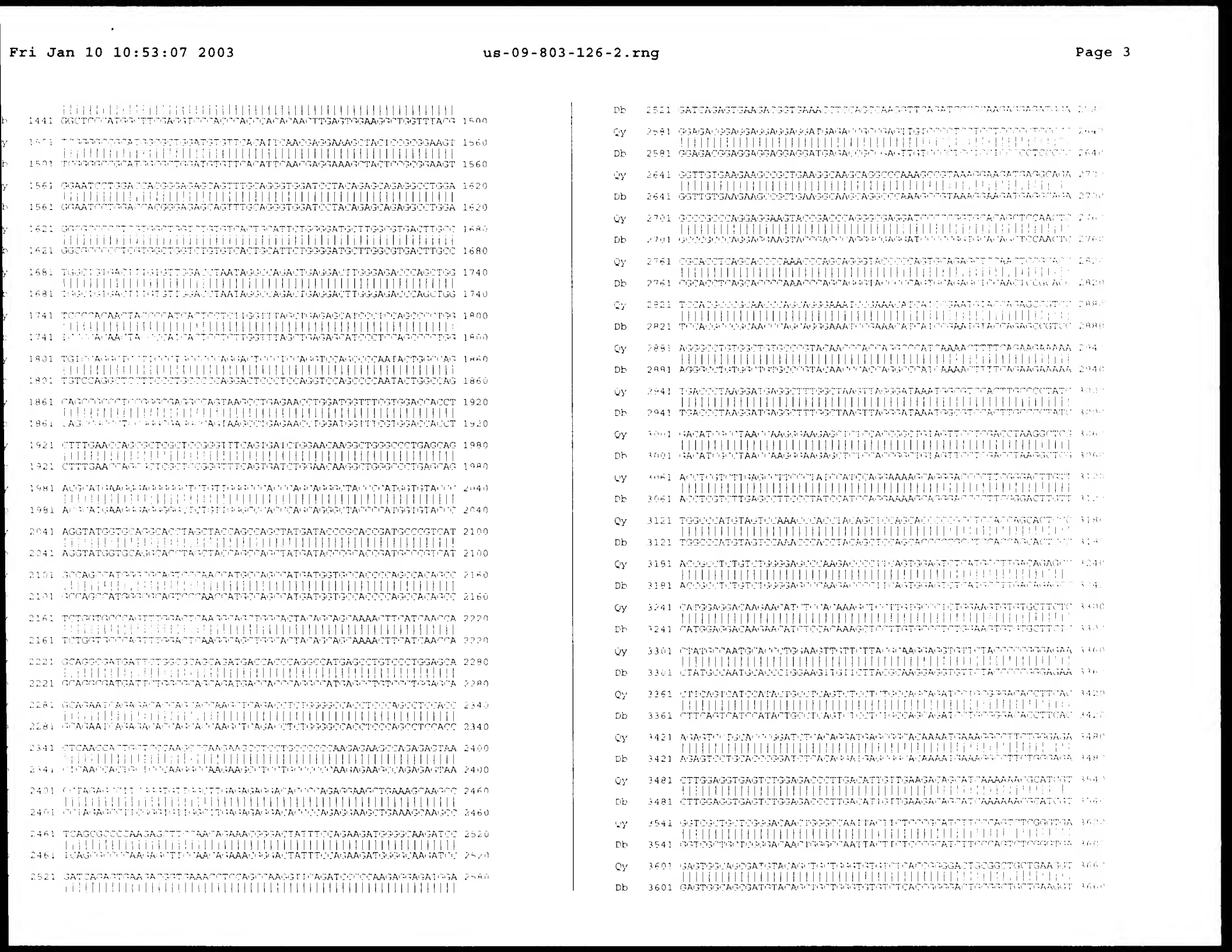


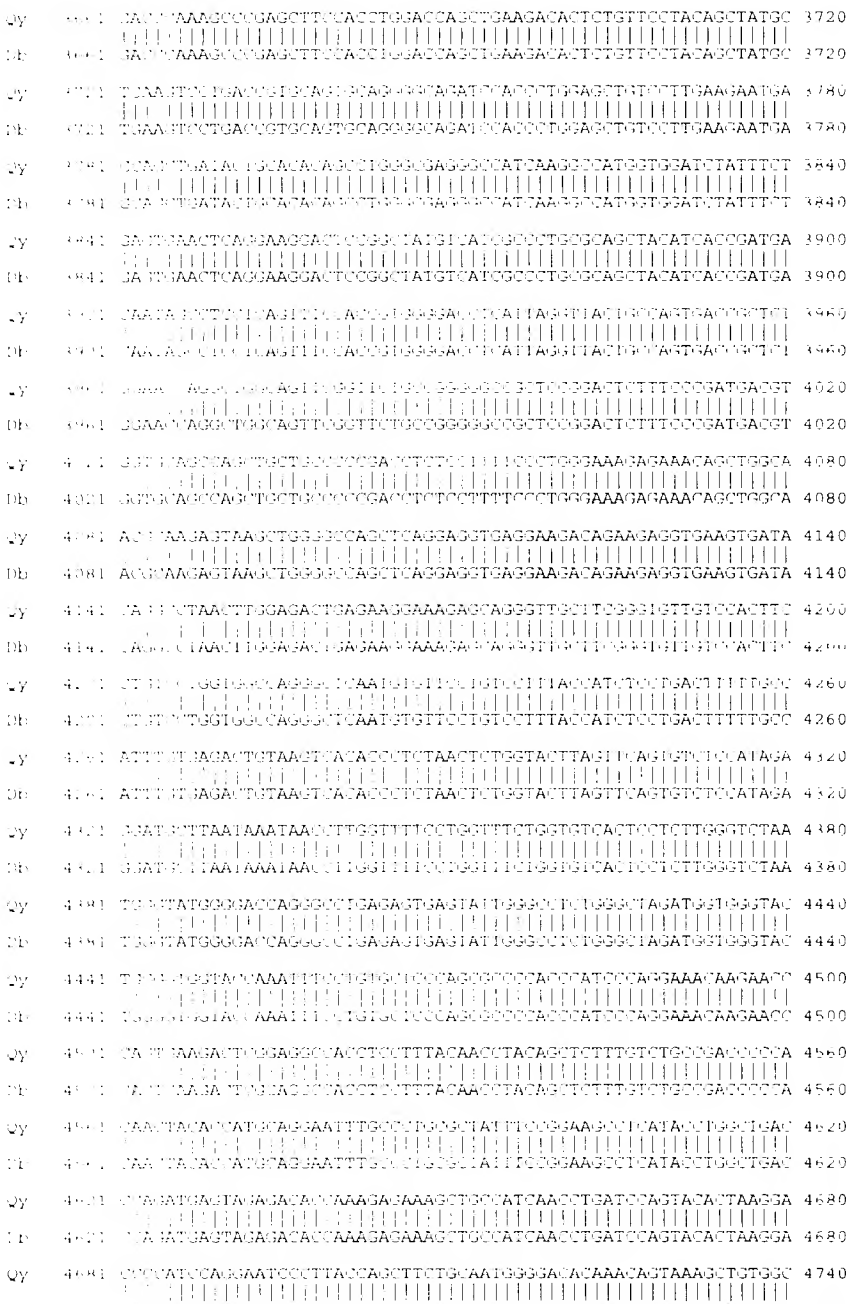
Search completed: January 9, 2003, 19:47:11

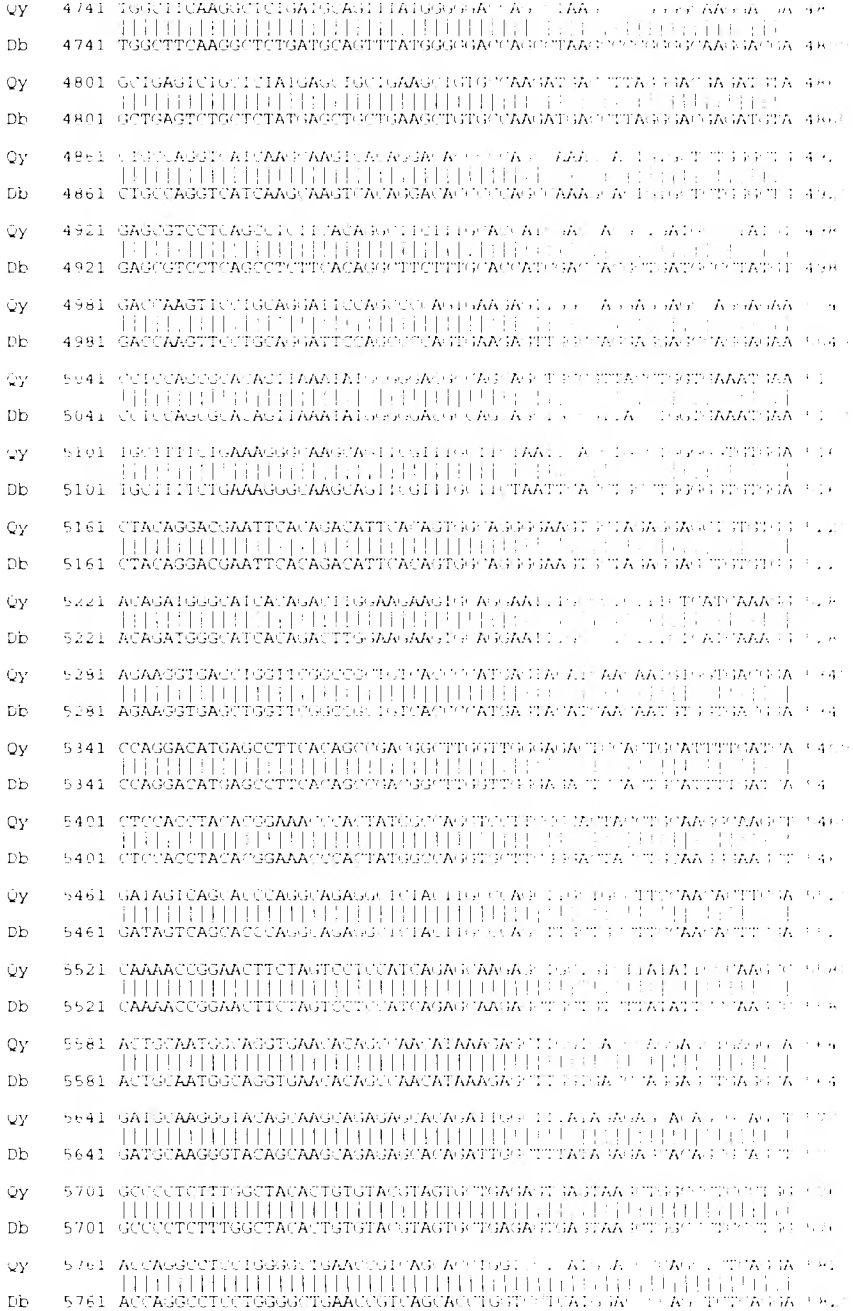
Job time : 15088 secs

```
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd

N nucleic - nucleic search, using sw model

un on:          January 9, 2003, 22:40:16 ; Search time 1034.91 Seconds
                                  (without alignments)
                                  1264.449 Million cell updates/sec

itle:           US-09-803-126-2
rfect score:    6293
equence:        1 [illegible] ......... aaaaaaaaaaaaaaaaaaa 6293

oring table:    IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

arched:         2185239 seqs, 1125999159 residues

tal number of hits satisfying chosen parameters:      437474

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing:  Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

tabase :        N_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/genesegn-embl/NA1988.DAT:*
                10: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1989.DAT:*
                11: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1990.DAT:*
                12: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1991.DAT:*
                13: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1992.DAT:*
                14: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1993.DAT:*
                15: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1994.DAT:*
                16: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1995.DAT:*
                17: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1996.DAT:*
                18: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1997.DAT:*
                19: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1998.DAT:*
                20: /SIDS2/gcgdata/geneseq/genesegn-embl/NA1999.DAT:*
                21: /SIDS2/gcgdata/geneseq/genesegn-embl/NA2000.DAT:*
                22: /SIDS2/gcgdata/geneseq/genesegn-embl/NA2001A.DAT:*
                23: /SIDS2/gcgdata/geneseq/genesegn-embl/NA2001B.DAT:*
                24: /SIDS2/gcgdata/geneseq/genesegn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
sult           Query
 No.   Score   Match Length DB  ID                Description
----------------------------------------------------------------------
   1   6293    100.0   6293 22  AAS13819          cDNA #1 encoding m
   2   4352     69.2   4375 22  AAS13820          cDNA #2 encoding m
   3   1787.6   28.4   4174 22  AAS13821          DNA encoding human
   4   1492     23.7   3780 22  AAS13823          DNA encoding human
   5    913     14.5   1960 24  ABK35127          Human cDNA encodin
   6    707.2   11.2   1346 22  AAS33274          DNA encoding human
   7    706.[illegible]  11.2   [illegible] 22  AAZ[illegible]  Human transmembran
   8    628.8   10.0   1044 22  AAD07791          Human secreted pro
   9    626.4   10.0   1268 23  ABK43544          DNA encoding novel
  10    621.2    9.9   1247 22  AAZ27110          cDNA encoding nove
  11    621.2    9.9   1247 23  ABK43853          DNA encoding novel
  12    470.2    7.5    926 22  AAD07777          Human secreted pro
c 13    166.2    2.6  24757 22  AAS[illegible]    DNA encoding novel
  14    166.2    2.6  24757 22  AAS33481          DNA encoding human
  15     97.6    1.6   [illegible] 22  AAS33480   DNA encoding human
c 16     94.4    1.5    566 22  ABA62139          Human foetal liver
c 17     94.4    1.5    566 22  ABA28517          Probe #8013 for ge
c 18     94.4    1.5    566 22  AAK10458          Human brain expres
c 19     94.4    1.5    566 22  AAK[illegible]    Human bone marrow
c 20     94.4    1.5    566 22  AAI42087          Probe #10773 used
c 21     94.4    1.5    566 24  ABS10413          Human genome deriv
c 22     92.2    1.5    233 22  ABA74633          Human foetal liver
c 23     92.2    1.5    233 22  ABA29397          Probe #[illegible] for g
c 24     92.2    1.5    233 22  AAK23124          Human brain expres
c 25     92.2    1.5    233 22  AAK49279          Human bone marrow
c 26     92.2    1.5    233 22  AAI55129          Probe #[illegible] used
c 27     92.2    1.5    233 24  ABS22813          Human genome deriv
c 28     89.4    1.4    526 22  ABA60390          Human foetal liver
c 29     89.4    1.4    526 22  ABA28621          Probe #7287 for ge
c 30     89.4    1.4    526 22  AAK08567          Human brain expres
c 31     89.4    1.4    526 22  AAK34556          Human bone marrow
c 32     89.4    1.4    526 22  AAI40274          Probe #8960 used
c 33     89.4    1.4    526 24  ABS09192          Human genome deriv
c 34     75.2    1.2    170 22  ABA72929          Human foetal liver
c 35     75.2    1.2    170 22  ABA28438          Probe #16964 for g
c 36     75.2    1.2    170 22  AAK21359          Human brain expres
c 37     75.2    1.2    170 22  AAK47502          Human bone marrow
c 38     75.2    1.2    170 22  AAI53353          Probe #22039 used
c 39     75.2    1.2    170 24  ABS21624          Human genome deriv
  40     65.4    1.0   5794 23  AAM78837          DNA encoding novel
  41     64.6    1.0   2794 24  ABN79632          Novel human coding
  42     64.2    1.0   1951 23  ABL04085          Drosophila melanog
  43     64.2    1.0   4966 23  ABL04081          Drosophila melanog
  44     61.4    1.0   1032 23  AAM78834          DNA encoding novel
  45     [illegible]  1.0   6733 23  ABL07105   Drosophila melanog

                              ALIGNMENTS

RESULT 1
AAS13819
ID   AAS13819 standard; cDNA; 6293 BP.
XX
AC   AAS13819;
XX
DT   19-DEC-2001  (first entry)
XX
DE   cDNA #1 encoding mouse myosin related protein, mMRP.
XX
KW   Mouse; oestrogen-regulated unconventional myosin related protein; MRP;
KW   nilitory; antiatherosclerotic; osteopathic; neuroprotective; nootropic;
KW   antiparkinsonian; cytostatic; oestrogen receptor; atherosclerosis;
KW   osteoporosis; breast cancer; cardiovascular disease; deafness; ss;
KW   Alzheimer's disease; Parkinson's disease; chromosome 11; immunogenic.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             125..4138
FT                   /*tag=  a
FT                   /product= "mMRP"
FT                   /note= "Oestrogen-regulated unconventional myosin
FT                          related protein"
XX
PN   WO200168666-A2.
XX
PD   20-SEP-2001.
XX
PF   12-MAR-2001; 2001WO-US08060.
XX
PR   10-MAR-2000; 2000US-188498P.
```

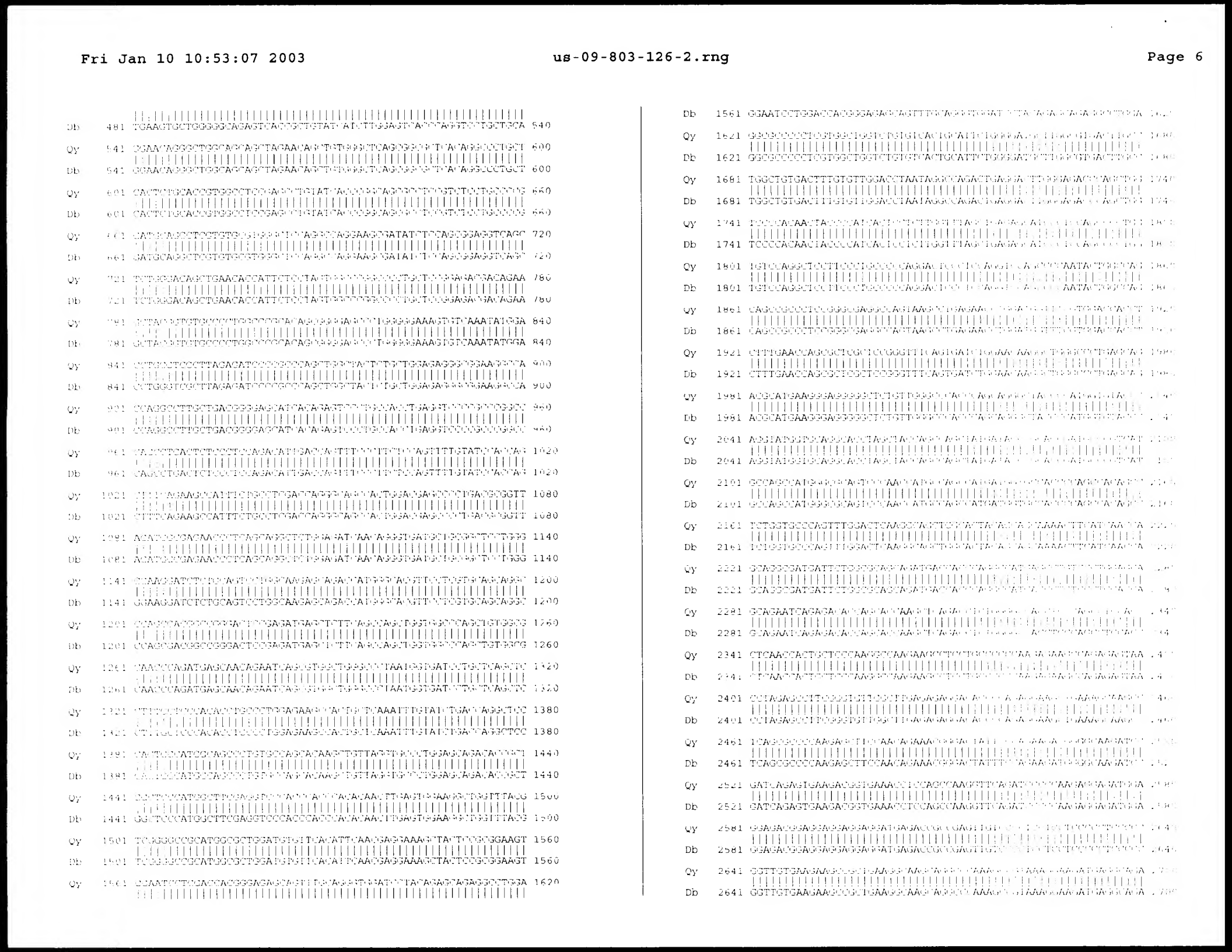
```
PR   [illegible] MAR 2001; 2001US 0803126.
XX
PA   (SCHE ) SCHERING AG.
XX
PI   Brooks AE,  Deng GG,  Rubanyi GM;
XX
DR   WPI; 2001 602747/68.
DR   P-PSDB; AAU07399.
XX
PT   Novel oestrogen-regulated unconventional myosin-related protein and
PT   polynucleotide for modulating effects of oestrogen in mammalian cells to
PT   treat, prevent atherosclerosis, breast cancer and osteoporosis -
XX
PS   Claim 4; Fig 2; 108pp; English.
XX
CC   The invention relates to an isolated oestrogen-regulated unconventional
CC   myosin related protein (MRP). Modulating MRP is useful for modulating the
CC   effects of oestrogen mediated by oestrogen receptor alpha or beta in a
CC   mammalian cell. When the level of expression or activity of MRP is
CC   increased, the development of atherosclerosis or osteoporosis is
CC   inhibited and when the level is decreased the development of breast
CC   cancer is inhibited in the mammal. Nucleic acid (I) encoding MRP is
CC   useful for detecting the presence of oestrogen signaling in a mammalian
CC   cell, by detecting the expression or activity of MRP or the level of MRP
CC   mRNA in the cell. The presence of oestrogen signaling is used to
CC   determine the responsiveness of the cell to the oestrogen or the tissue-
CC   specific distribution of oestrogen signaling in a mammal. MRP is useful
CC   as a marker gene for the investigation of tissue-specific and oestrogen
CC   receptor specific ligands such as oestrogen receptor antagonists and
CC   agonists. Modulators, recombinant forms, derivatives, variants or
CC   fragments of MRP are useful for enhancing or inhibiting oestrogen
CC   signaling in cells and for treatment of breast cancer or other cancers
CC   and for preventing or treating wide variety of oestrogen-related
CC   diseases, including osteoporosis, cardiovascular diseases, Alzheimer's
CC   disease or Parkinson's disease. Modulation of MRP gene is also useful for
CC   preventing and/or treating hearing loss and deafness. MRP polypeptides
CC   are useful for generating monoclonal and polyclonal antibodies useful for
CC   identifying oestrogen responsive cells, particularly cells responsive to
CC   oestrogen through an oestrogen receptor alpha or beta, to identify
CC   alleles, interspecies homologues and polymorphic variants and in
CC   immunoassays to detect MRP. The present sequence represents the coding
CC   sequence of mouse MRP variant #1 as described in the invention.
XX
SQ   Sequence 6293 BP; 1433 A; 1917 C; 1733 G; 1210 T; 0 other;

  Query Match             100.0%;  Score 6293;  DB 22;  Length 6293;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 6293;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 [illegible]AGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60
Db          1 [illegible]AGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGGGGCCACCTGGA 60

Qy         61 [illegible]GCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120
Db         61 [illegible]GCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120

Qy        121 [illegible]ATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGGCTGGGACTGAGCAAAACAAACC 180
Db        121 [illegible]ATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGGCTGGGACTGAGCAAAACAAACC 180

Qy        181 CACATTGGCCTCTCGATTCCAGCAGACCCTGGGTGACTTGCTAGCTGGGCTAGGCAGCAG 240
Db        181 CACATTGGCCTCTCGATTCCAGCAGACCCTGGGTGACTTGCTAGCTGGGCTAGGCAGCAG 240

Qy        241 [illegible]CAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300
Db        241 [illegible]CAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300

Qy        301 [illegible]GCAGAGCAGCTGCGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360
Db        301 [illegible]GCAGAGCAGCTGCGTCAGGCTGGCATCCTGGAGATCATAGGCAC 360

Qy        361 CCGGAGTACCCACTTC[illegible] 420
Db        361 CCGGAGTACCCACTTC[illegible] 420

Qy        421 CCTGGGGTCAGGGAGACAGAAAGC[illegible] 480
Db        421 CCTGGGGTCAGGGAGACAGAAAGC[illegible] 480

Qy        481 TGAAGT[illegible] 540
Db        481 TGAAGT[illegible] 540

Qy        541 GGAACAGGG[illegible]AGCTAGAACAG[illegible] 600
Db        541 GGAACAGGG[illegible]AGCTAGAACAG[illegible] 600

Qy        601 [illegible] 660
Db        601 [illegible] 660

Qy        661 GATGCAGGCTCGTGTGCGTGGGCTCCAG[illegible] 720
Db        661 GATGCAGGCTCGTGTGCGTGGGCTCCAG[illegible] 720

Qy        721 [illegible] 780
Db        721 [illegible] 780

Qy        781 [illegible]AAA[illegible]TCAAATAT[illegible] 840
Db        781 [illegible]AAA[illegible]TCAAATAT[illegible] 840

Qy        841 [illegible] 900
Db        841 [illegible] 900

Qy        901 CCAGGCCTTGCTGACGGGGAGCAT[illegible] 960
Db        901 CCAGGCCTTGCTGACGGGGAGCAT[illegible] 960

Qy        961 CAGCCTGACT[illegible] 1020
Db        961 CAGCCTGACT[illegible] 1020

Qy       1021 CTTTCAGAAG[illegible] 1080
Db       1021 CTTTCAGAAG[illegible] 1080

Qy       1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAA[illegible] 1140
Db       1081 AGATGGCGAGAACCCTCAGCAGGCTCTGGAGATCAA[illegible] 1140

Qy       1141 GGAAGGATC[illegible] 1200
Db       1141 GGAAGGATC[illegible] 1200

Qy       1201 CCAGGGACGG[illegible] 1260
Db       1201 CCAGGGACGG[illegible] 1260

Qy       1261 CAACCCAGATGAGCAACAGAATCAG[illegible] 1320
Db       1261 CAACCCAGATGAGCAACAGAATCAG[illegible] 1320

Qy       1321 CTTTGCTCCCACACCTGCCCTGGAGAAG[illegible] 1380
Db       1321 CTTTGCTCCCACACCTGCCCTGGAGAAG[illegible] 1380

Qy       1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAG[illegible] 1440
Db       1381 CAGTGGCATGGCAGCCCTGTGCCAGCACAAG[illegible] 1440

Qy       1441 [illegible] 1500
```

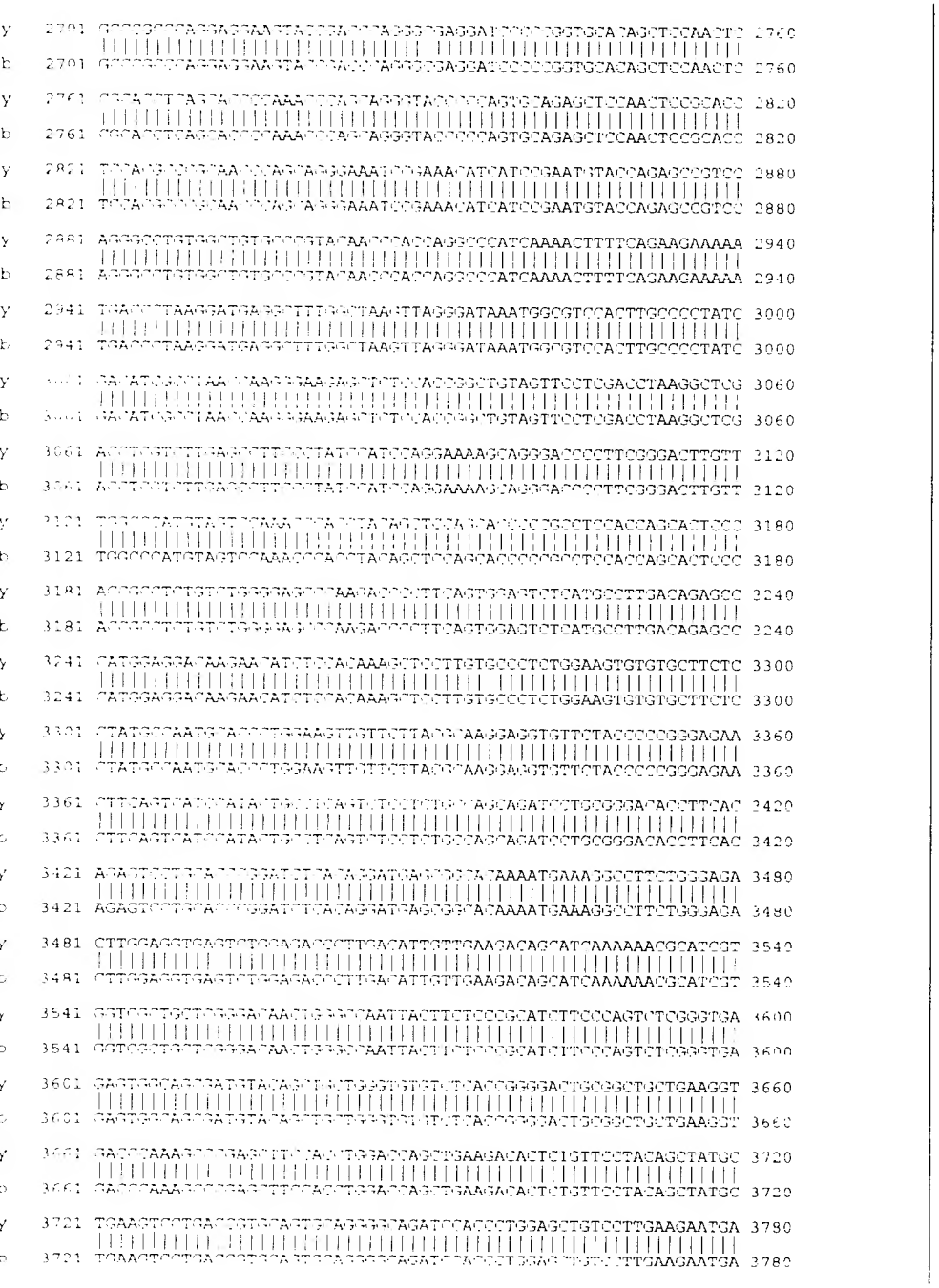






```
Db    5871 ACTCTGCTGCTCTGTCATGCTAAAAGACCTGAAGCAGTTCCACCTGTTGAGCCCACTGCA 5930

Qy    5881 GGAGGACGGGCCCCCTGGCCTAGAACTCAACTATGGCTCTGTTGACAACCCCCAGACCAT 5940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5881 GGAGGACGGGCCCCCTGGCCTAGAACTCAACTATGGCTCTGTTGACAACCCCCAGACCAT 5940

Qy    5941 CTGGTTGGAGTTGCCACAGGCCAGGAGCTGCAGCACACCATCATCTTCCTGCTGGGCAG 6000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5941 CTGGTTGGAGTTGCCACAGGCCAGGAGCTGCAGCACACCATCATCTTCCTGCTGGGCAG 6000

Qy    6001 CATGTCCACTCAGTGGCCAGGTCTCCTGAGGAAGTCAAGATAAGGTACGGTCTCTCAC 6060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6001 CATGTCCACTCAGTGGCCAGGTCTCCTGAGGAAGTCAAGATAAGGTACGGTCTCTCAC 6060

Qy    6061 TGGGCAGTTGGTTTAGTCCTGCTCGAAAGCGCTGCACAACCCCCACCCCACGTGGAG 6120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6061 TGGGCAGTTGGTTTAGTCCTGCTCGAAAGCGCTGCACAACCCCCACCCCACGTGGAG 6120

Qy    6121 GCCAAAAGGCAAAGTTGTGTCACCTGGGAGAATAGGCAGACACATCCCCTCTGGGGTGGA 6180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6121 GCCAAAAGGCAAAGTTGTGTCACCTGGGAGAATAGGCAGACACATCCCCTCTGGGGTGGA 6180

Qy    6181 CTGCAACAGGAGTTGGGGCATTTGCTGGCTAGCCCCAGGGAAAATGCCCACCCAGCTCGA 6240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6181 CTGCAACAGGAGTTGGGGCATTTGCTGGCTAGCCCCAGGGAAAATGCCCACCCAGCTCGA 6240

Qy    6241 AAGCGGCACAAGTAAAACACCCAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6241 AAGCGGCACAAGTAAAACACCCAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6297


ESULT 2
AS13820
D   AAS13820 standard; cDNA; 4375 BP.
X
C   AAS13820;
X
T   18-DEC-2001 (first entry)
X
E   cDNA #2 encoding mouse myosin related protein, mMRP.
X
W   Mouse; oestrogen regulated unconventional myosin-related protein, MRP,
W   auditory; antiatherosclerotic; osteopathic; neuroprotective; nootropic;
W   antiparkinsonian; cytostatic; oestrogen receptor; atherosclerosis;
W   osteoporosis, breast cancer; cardiovascular disease; deafness; ss;
W   Alzheimer's disease, Parkinson's disease, chromosome 11; immunogenic
X
S   Mus sp.
X
H   Key             Location/Qualifiers
T   CDS             125..4138
T                   /*tag= a
T                   /product= "mMRP"
T                   /note= "Oestrogen-regulated unconventional myosin-
T                         related protein"
X
N   WO200168866-A2.
X
D   20-SEP-2001.
X
F   12-MAR-2001; 2001WO-US08060.
X
R   10-MAR-2000; 2000US-188488P.
R   09-MAR-2001; 2001US-0803126.
X
A   (SCHD ) SCHERING AG.
X
I   Brooks AR, Peng SG, Elkhanyi GM.
X
R   WPI; 2001-603747/69.
R   P-PSDB; AAU07399.
X
PT  Novel oestrogen-regulated unconventional myosin-related protein and
PT  polynucleotide for modulating effects of oestrogen in mammalian cells to
PT  treat, prevent atherosclerosis, breast cancer and osteoporosis
XX
PS  Claim 4; Fig 3; 108pp; English.
XX
CC  The invention relates to an isolated oestrogen-regulated unconventional
CC  myosin-related protein (MRP). Modulating MRP is useful for modulating the
CC  effects of oestrogen mediated by oestrogen receptor alpha or beta in a
CC  mammalian cell. When the level of expression or activity of MRP is
CC  increased, the development of atherosclerosis or osteoporosis is
CC  inhibited and when the level is decreased the development of breast
CC  cancer is inhibited in the mammal. Nucleic acid (I) encoding MRP is
CC  useful for detecting the presence of oestrogen signaling in a mammalian
CC  cell, by detecting the expression or activity of MRP or the level of MRP
CC  mRNA in the cell. The presence of oestrogen signaling is used to
CC  determine the responsiveness of the cell to the oestrogen or the tissue
CC  specific distribution of oestrogen signaling in a mammal. MRP is useful
CC  as a marker gene for the investigation of tissue-specific and oestrogen
CC  receptor-specific ligands such as oestrogen receptor antagonists and
CC  agonists. Modulators, recombinant forms, derivatives, variants or
CC  fragments of MRP are useful for enhancing or inhibiting oestrogen
CC  signaling in cells and for treatment of breast cancer or other cancers
CC  and for preventing or treating wide variety of oestrogen related
CC  diseases, including osteoporosis, cardiovascular diseases, Alzheimer's
CC  disease or Parkinson's disease. Modulation of MRP gene is also useful for
CC  preventing and/or treating hearing loss and deafness. MRP polypeptides
CC  are useful for generating monoclonal and polyclonal antibodies useful for
CC  identifying oestrogen responsive cells, particularly cells responsive to
CC  oestrogen through an oestrogen receptor alpha or beta, to identify
CC  alleles, interspecies homologues and polymorphic variants and in
CC  immunoassays to detect MRP. The present sequence represents the coding
CC  sequence of mouse MRP variant #2 as described in the invention.
XX
SQ  Sequence 4375 BP; 966 A, 1368 C, 1116 G, 915 T, 10 other;

  Query Match             69.2%;  Score 4352;  DB 22;  Length 4375;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4352;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CGCTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGAGGCCACCTGGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGCTGGGACTGTCACCTACCAGGTGCACAAGTTCATAAACAGAAACAGAGGCCACCTGGA 60

Qy     61 CCCCGCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCCCGCTGTGCTGGAGATGCTCAGGCAGAGCCAGCTGCAGGTGACCTAGCCTTCCTTTCA 120

Qy    121 GCTCATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGCTTGGGACTGAGCAAAACAAACC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTCATGGGCAGCCTGTTCCAAGAAGCAGAGCCCCAGCTTGGGACTGAGCAAAACAAACC 180

Qy    181 CACATTGGCCTCTGGATTCCAGCAGACCTGGGTGACTTGCTAGCTCCGCTAGGCAGCAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CACATTGGCCTCTGGATTCCAGCAGACCTGGGTGACTTGCTAGCTCCGCTAGGCAGCAG 240

Qy    241 GGGCCATGTCTACGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGCCATGTCTACGTCATCCACTGTCTCAATCCCACCCCTGGAAAGATCCCAGGCCTCTT 300

Qy    301 GGAGTGGGGCATGTGGCAGAGCAGCTGGTCAGCTGGCATCCTGGAGATCATAGGCAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAGTGGGGCATGTGGCAGAGCAGCTGGTCAGCTGGCATCCTGGAGATCATAGGCAC 360

Qy    361 CGGGAGTACCCACTTCCCCGTGCGAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGGGAGTACCCACTTCCCCGTGCGAGTGTCCTTCCAAGTCTTTCTGGCAAGGTTCCATGC 420

Qy    421 CCTGGGGTCAGGGAGACAGAAAGTGCCTCTGACCAGGAGAGGTGTGGTGCCAGCCTCAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTGGGGTCAGGGAGACAGAAAGTGCCTCTGACCAGGAGAGGTGTGGTGCCAGCCTCAG 480

Qy    481 TGAAGTGCTGGGGGCAGAGTCACCGCTGTATCATCTTGGAGTCACCCAGGTCCTGCTGCA 540
```



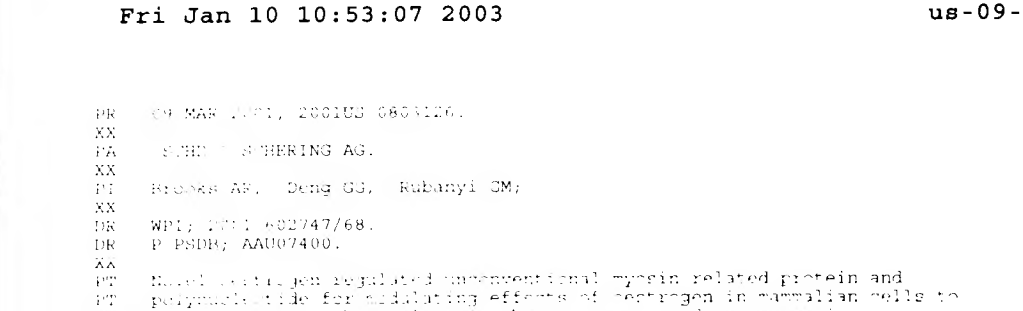


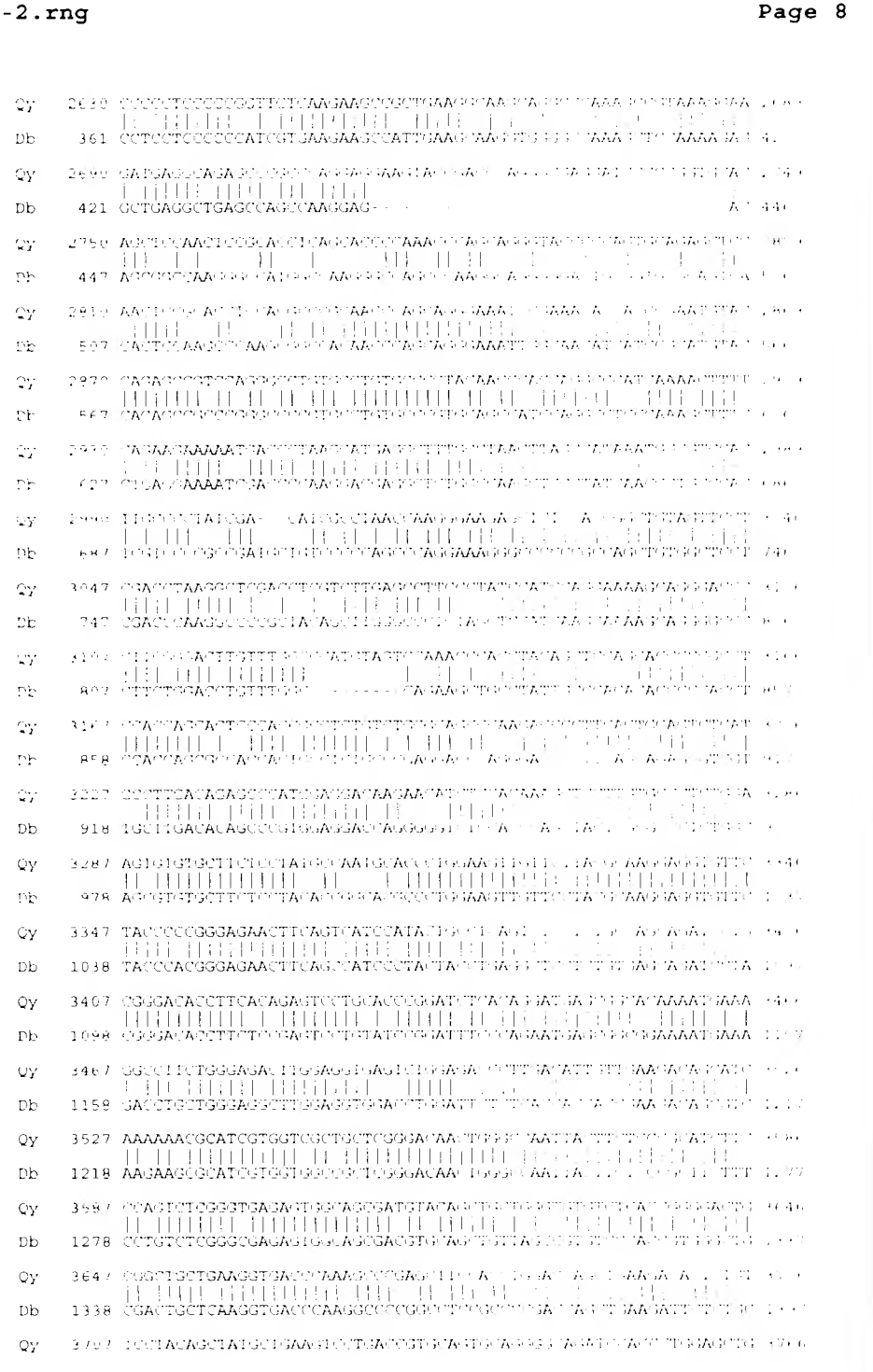
```
Qy   3781 GCAGCTGATACTGCACACAGCCTGGGCGAGGGCCATCAAGGCCATGGTGGATCTATTTCT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GCAGCTGATACTGCACACAGCCTGGGCGAGGGCCATCAAGGCCATGGTGGATCTATTTCT 3840

Qy   3841 GAGTGAACTCAGGAAGGACTCCGGCTATGTCATCGCCCTGGCGAGCTACATCACCGATGA 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 GAGTGAACTCAGGAAGGACTCCGGCTATGTCATCGCCCTGGCGAGCTACATCACCGATGA 3900

Qy   3901 CAATAGCCTCCTCAGTTTCCACGGTGGGGACCTCATTAGGTTACTGCCAGTGACCGCTC1 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CAATAGCCTCCTCAGTTTCCACGGTGGGGACCTCATTAGGTTACTGCCAGTGACCGCTC1 3960

Qy   3961 GGAACCAGGCTGGCAGTTCGGTTCTGCCGGGGGCCGCTCTGGGACTCTTTCCCGATGACGT 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GGAACCAGGCTGGCAGTTCGGTTCTGCCGGGGGCCGCTCTGGGACTCTTTCCCGATGACGT 4020

Qy   4021 GGTGCAGCCAGCTGCTGCCCCCCGACCTCTCCTTTTCCCTGGGAAAGAGAAACAGCTGGCA 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GGTGCAGCCAGCTGCTGCCCCCGACCTCTCCTTTTCCCTGGGAAAGAGAAACAGCTGGCA 4080

Qy   4081 ACGCAAGAGTAAGCTGGGGCCAGCTCAGGAGGTGAGGAAGACAGAAGAGGTGAAGTGATA 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 ACGCAAGAGTAAGCTGGGGCCAGCTCAGGAGGTGAGGAAGACAGAAGAGGTGAAGTGATA 4140

Qy   4141 CAGGCCTAACTTGGAGACTGAGAAGGAAAGAGCAGGGTTGCTTCGGGTGTGTGCCACTTC 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 CAGGCCTAACTTGGAGACTGAGAAGGAAAGAGCAGGGTTGCTTCGGGTGTGTGCCACTTC 4200

Qy   4201 CTGTCCTGGTGGCCAGGGCTCAATGTGTTCCTGTCCTTTACCATCTCCTGACTTTTTGCC 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CTGTCCTGGTGGCCAGGGCTCAATGTGTTCCTGTCCTTTACCATCTCCTGACTTTTTGCC 4260

Qy   4261 ATTTGTGAGACTGTAAGTCACACCCTCTAACTCTGGTACTTAGTTCAGTGTCTCCATAGA 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 ATTTGTGAGACTGTAAGTCACACCCTCTAACTCTGGTACTTAGTTCAGTGTCTCCATAGA 4320

Qy   4321 GGATGCTTAATAAATAACCTTGGTTTTCCTGG 4352
          ||||||||||||||||||||||||||||||||
Db   4321 GGATGCTTAATAAATAACCTTGGTTTTCCTGG 4352
```

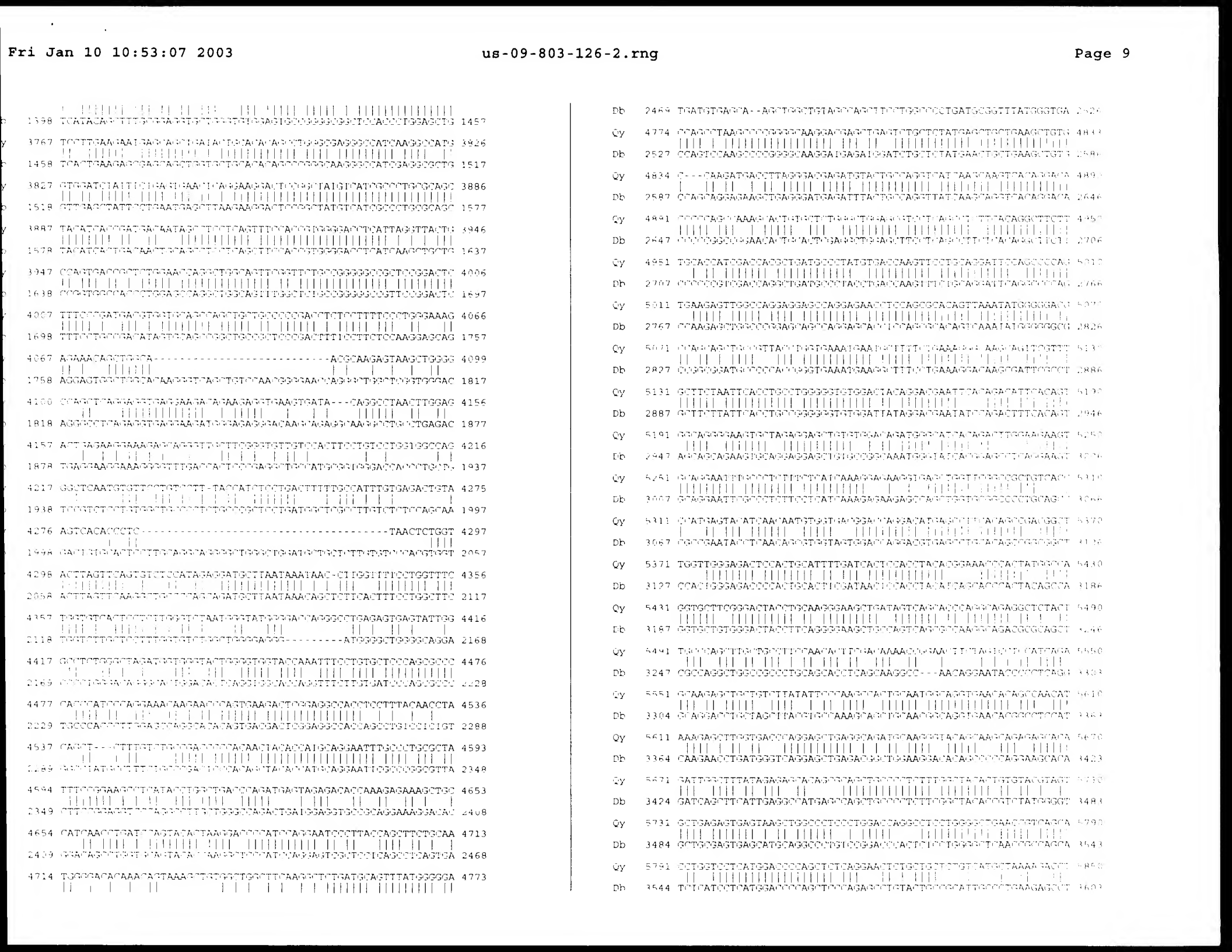
RESULT 3
AAS13821
ID AAS13821 standard, cDNA, 4174 BP.
XX
AC AAS13821;
XX
DT 18-DEC-2001 (first entry)
XX
DE DNA encoding human myosin-related protein, hMRP #1.
XX
KW Human; oestrogen regulated unconventional myosin-related protein; MRP;
KW auditory; antiatherosclerotic, osteopathic; neuroprotective, neotropic;
KW antiparkinsonian, cytostatic; oestrogen receptor, atherosclerosis;
KW osteoporosis; breast cancer; cardiovascular disease; deafness; ss;
KW Alzheimer's disease, Parkinson's disease; chromosome 17q25, immunogenic.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   561..1874
FT   /*tag= a
FT   /product= "hMRP"
FT   /note= "Oestrogen regulated unconventional myosin
FT   related protein"
XX
PN WO200168866-A2.
XX
PD 20-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-US08060.
XX
PR 10-MAR-2000; 2000US-188488P.

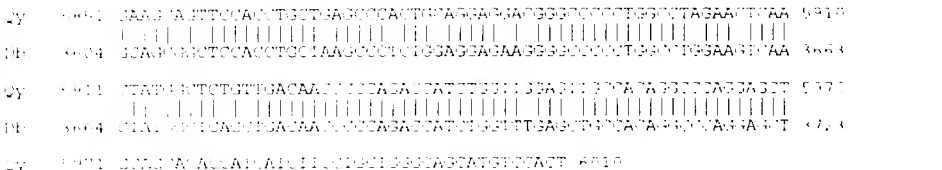
PR [illegible] MAR 2001; 2001US-0803126.
XX
PA [illegible] SCHERING AG.
XX
PI Brooks AR, Deng GG, Rubanyi GM;
XX
DR WPI; 2001-602747/68.
DR P-PSDB; AAU07400.
XX
PT Novel oestrogen regulated unconventional myosin related protein and
PT polynucleotide for modulating effects of oestrogen in mammalian cells to
PT treat, prevent atherosclerosis, breast cancer and osteoporosis -
XX
PS Claim 4; Fig 5; 108pp; English.
XX
CC The invention relates to an isolated oestrogen-regulated unconventional
CC myosin related protein (MRP). Modulating MRP is useful for modulating the
CC effects of oestrogen mediated by oestrogen receptor alpha or beta in a
CC mammalian cell. When the level of expression or activity of MRP is
CC increased, the development of atherosclerosis or osteoporosis is
CC inhibited and when the level is decreased the development of breast
CC cancer is inhibited in the mammal. Nucleic acid (I) encoding MRP is
CC useful for detecting the presence of oestrogen signaling in a mammalian
CC cell, by detecting the expression or activity of MRP or the level of MRP
CC mRNA in the cell. The presence of oestrogen signaling is used to
CC determine the responsiveness of the cell to the oestrogen or the tissue-
CC specific distribution of oestrogen signaling in a mammal. MRP is useful
CC as a marker gene for the investigation of tissue-specific and oestrogen
CC receptor specific ligands such as oestrogen receptor antagonists and
CC agonists. Modulators, recombinant forms, derivatives, variants or
CC fragments of MRP are useful for enhancing or inhibiting oestrogen
CC signaling in cells and for treatment of breast cancer or other cancers
CC and for preventing or treating wide variety of oestrogen related
CC diseases, including osteoporosis, cardiovascular diseases, Alzheimer's
CC disease or Parkinson's disease. Modulation of MRP gene is also useful for
CC preventing and/or treating hearing loss and deafness. MRP polypeptides
CC are useful for generating monoclonal and polyclonal antibodies useful for
CC identifying oestrogen responsive cells, particularly cells responsive to
CC oestrogen through an oestrogen receptor alpha or beta, to identify
CC alleles, interspecies homologues and polymorphic variants and in
CC immunoassays to detect MRP. The present sequence represents the coding
CC sequence of human MRP variant #1 as described in the invention.
XX
SQ Sequence 4174 BP; 901 A; 1339 C; 1237 G; 697 T; 0 other;

Query Match 28.4%; Score 1787.6; DB 22; Length 4174;
Best Local Similarity 72.3%; Pred No. 0;
Matches 1[illegible]; Conservative 2; Mismatches 969; Indels 166; Gaps 15;

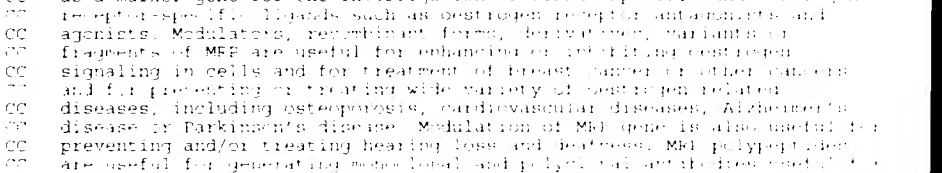




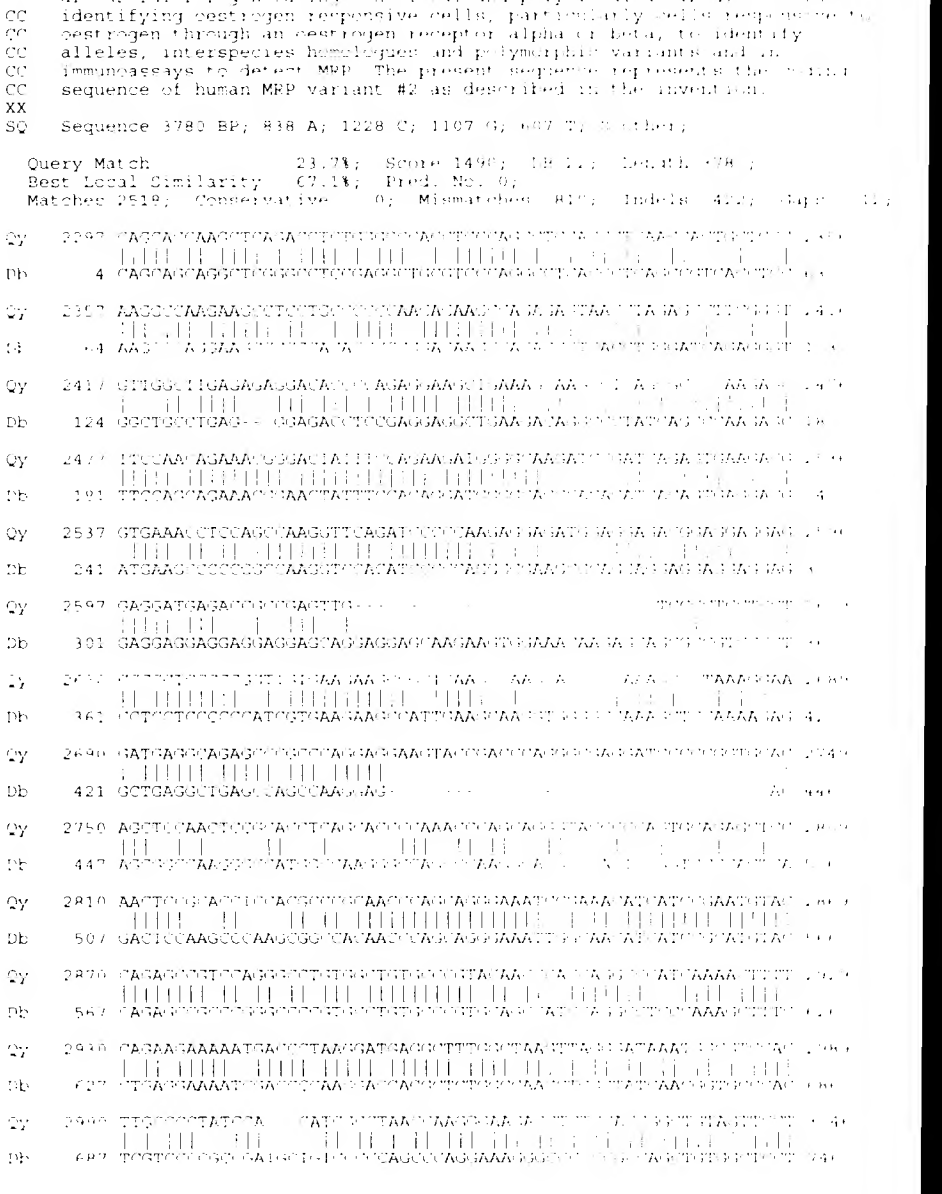


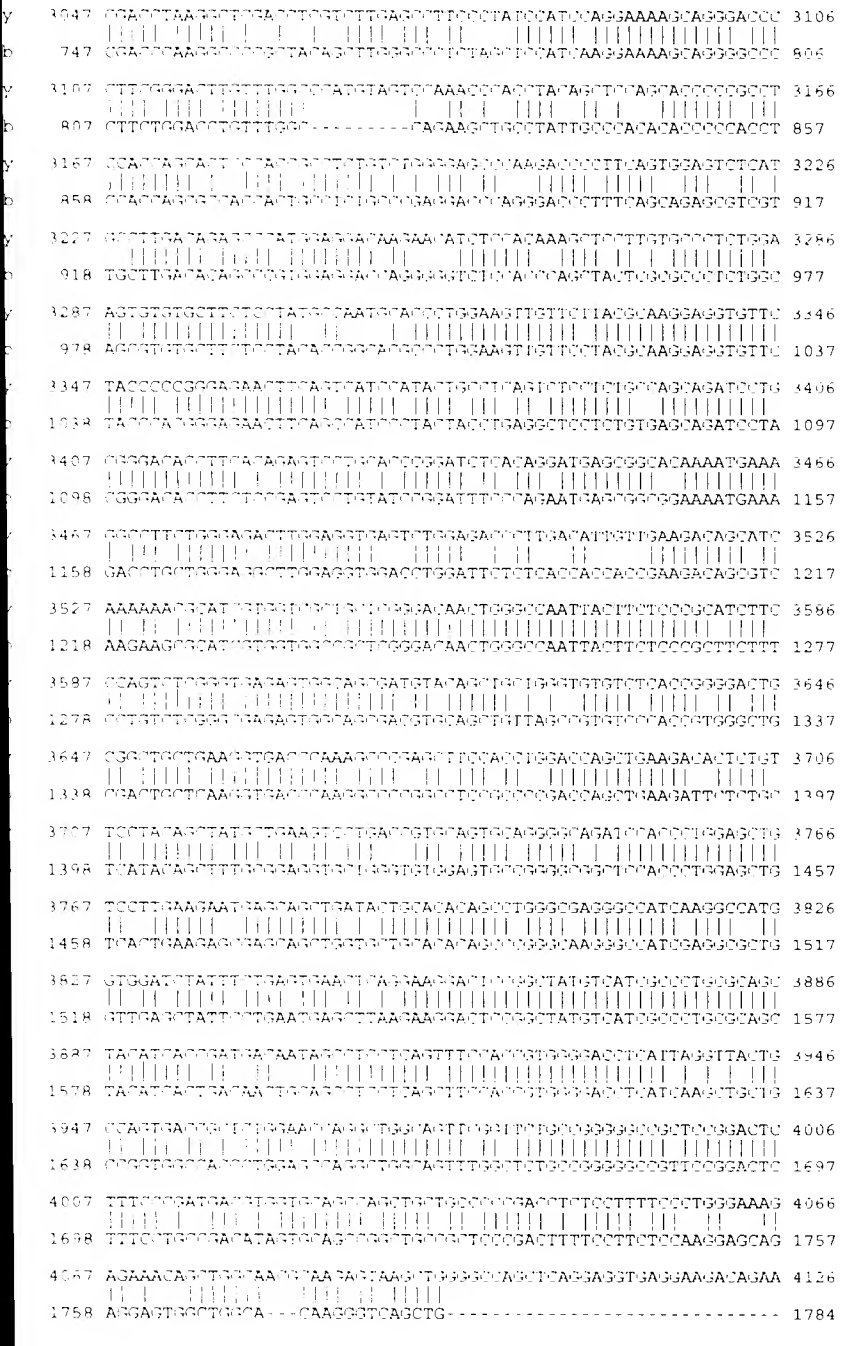


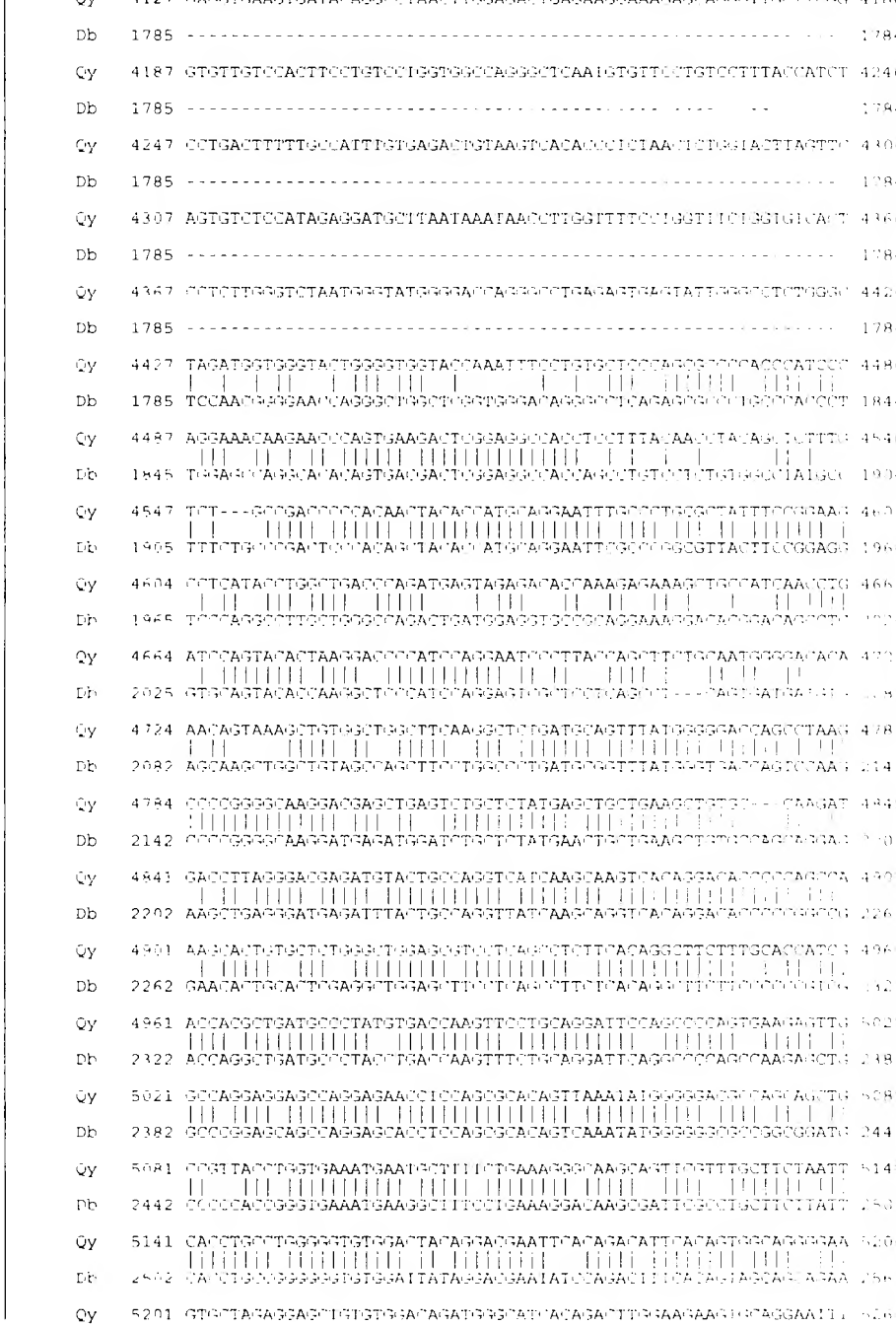
RESULT 4
AAS13823
ID AAS13823 standard; cDNA; 3780 BP.
XX
AC AAS13823;
XX
DT 19-DEC-2001 (first entry)
XX
DE DNA encoding human myosin related protein, hMRP #2
XX
KW Human; oestrogen regulated unconventional myosin-related protein; MRP;
KW auditory; antiatherosclerotic; osteopathic; neuroprotective; nootropic;
KW antiproliferative, cytostatic, oestrogen receptor, atherosclerosis;
KW osteoporosis; breast cancer; cardiovascular disease; deafness; ss;
KW Alzheimer's disease, Parkinson's disease, chromosome 17p13, immunogenic
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 561..2911
FT /*tag= a
FT /product= "hMRP"
FT /note= "Oestrogen regulated unconventional myosin-
FT related protein"
XX
PN WO200168868-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-US08060.
XX
PR 16-MAR-2000; 2000US-188488P.
PR 09-MAR-2001; 2001US-0803126.
XX
PA (SCHE) SCHERING AG.
XX
PI Fricke AD, Dena CG, Rubanyi GM;
XX
DR WPI; 2001-602747/68.
DR P-PSDB; AAU07406.
XX
PT Novel oestrogen regulated unconventional myosin related protein and
PT polynucleotide for modulating effects of oestrogen in mammalian cells to
PT treat, prevent atherosclerosis, breast cancer and osteoporosis
XX
PS Claim 4; Fig 7; 108pp; English.
XX
CC The invention relates to an isolated oestrogen-regulated unconventional
CC myosin related protein (MRP). Modulating MRP is useful for modulating the
CC effects of oestrogen mediated by oestrogen receptor alpha or beta in a
CC mammalian cell. When the level of expression or activity of MRP is
CC increased, the development of atherosclerosis or osteoporosis is
CC inhibited and when the level is decreased the development of breast
CC cancer is inhibited in the mammal. Nucleic acid (I) encoding MRP is
CC useful for detecting the presence of oestrogen signaling in a mammalian
CC cell, by detecting the expression or activity of MRP or the level of MRP
CC mRNA in the cell. The presence of oestrogen signaling is used to
CC determine the responsiveness of the cell to the oestrogen or the tissue-
CC specific distribution of oestrogen signaling in a mammal. MRP is useful
CC as a marker gene for the investigation of tissue specific and cell type
CC receptor-specific ligands such as oestrogen receptor antagonists and
CC agonists. Modulators, recombinant forms, derivatives, variants or
CC fragments of MRP are useful for enhancing or inhibiting oestrogen
CC signaling in cells and for treatment of breast cancer or other cancers
CC and for preventing or treating a wide variety of oestrogen related
CC diseases, including osteoporosis, cardiovascular diseases, Alzheimer's
CC disease or Parkinson's disease. Modulation of MRP gene is also useful for
CC preventing and/or treating hearing loss and deafness. MRP polypeptides
CC are useful for generating monoclonal and polyclonal antibodies used for
CC identifying oestrogen responsive cells, particularly cells responsive to
CC oestrogen through an oestrogen receptor alpha or beta, to identify
CC alleles, interspecies homologues and polymorphic variants and in
CC immunoassays to detect MRP. The present sequence represents the nucleic
CC sequence of human MRP variant #2 as described in the invention.
XX
SQ Sequence 3780 BP; 838 A; 1228 C; 1107 G; 607 T; 0 other;

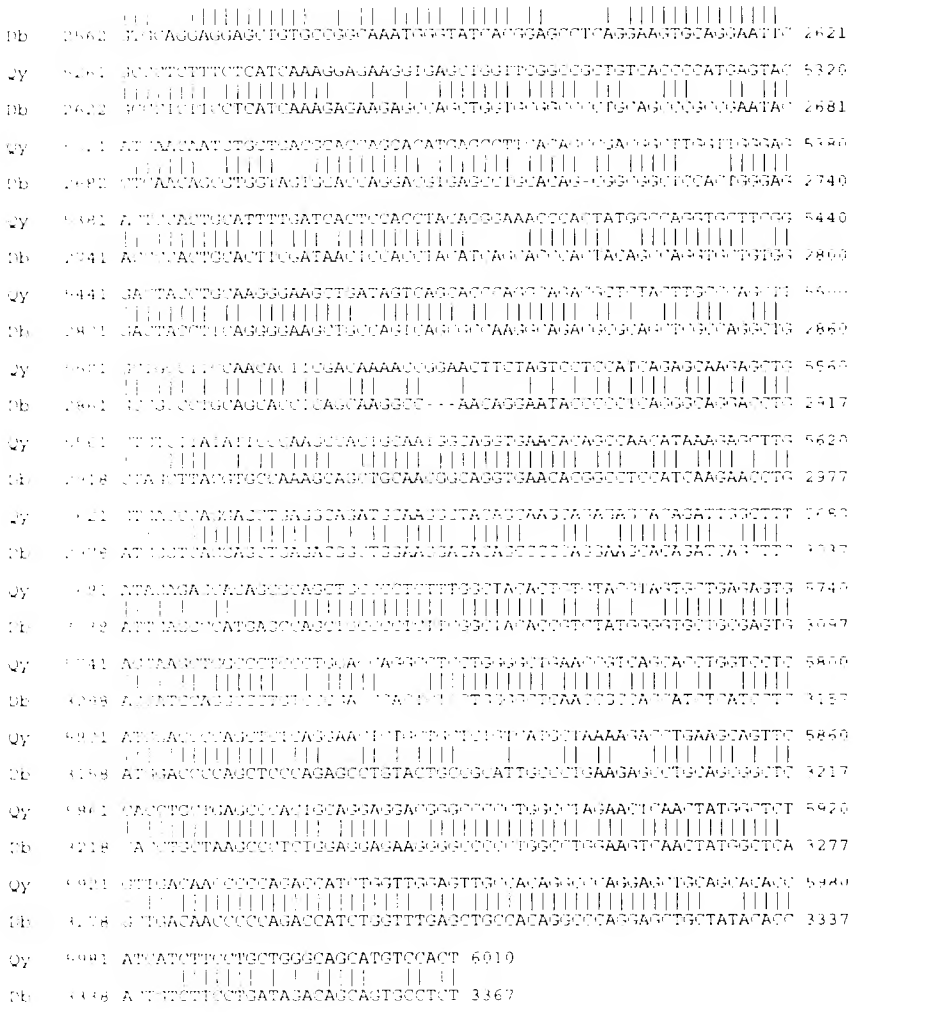


Query Match 23.7%; Score 1490; 18.1; Length 378;
Best Local Similarity 67.1%; Prod. No. 0;
Matches 2519; Conservative 0; Mismatches 811; Indels 422; Gaps 11;

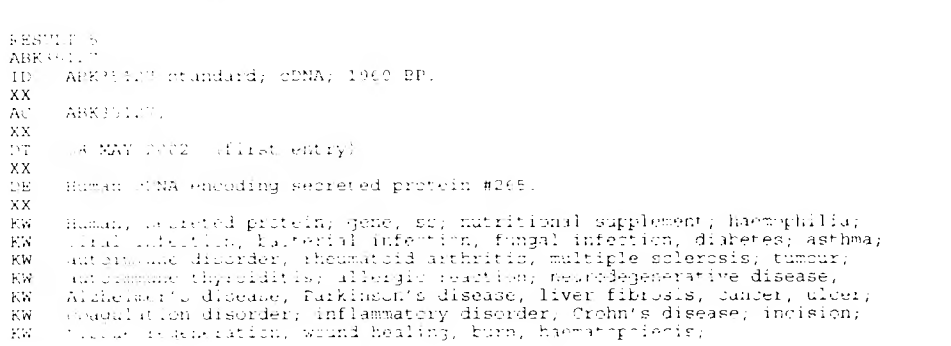




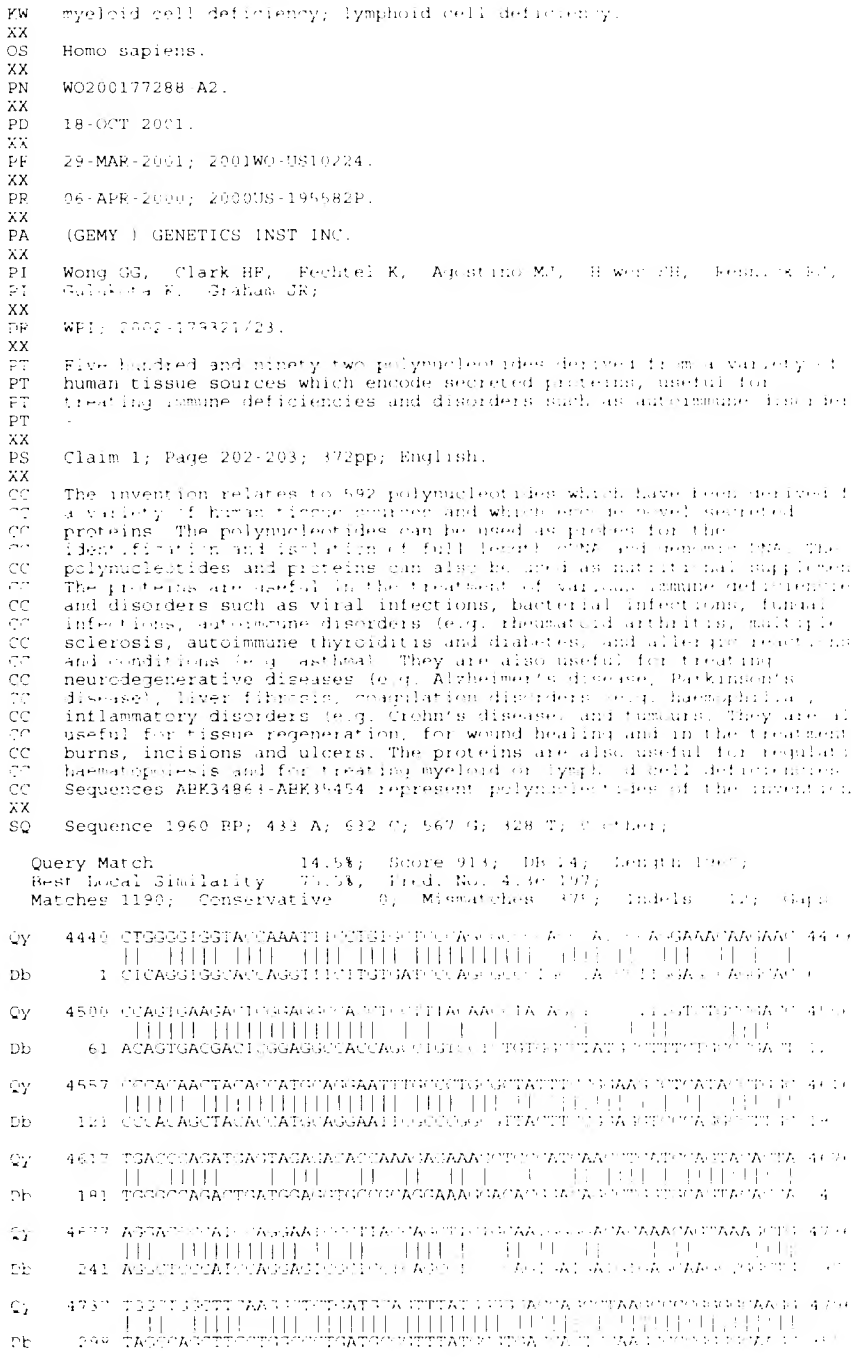


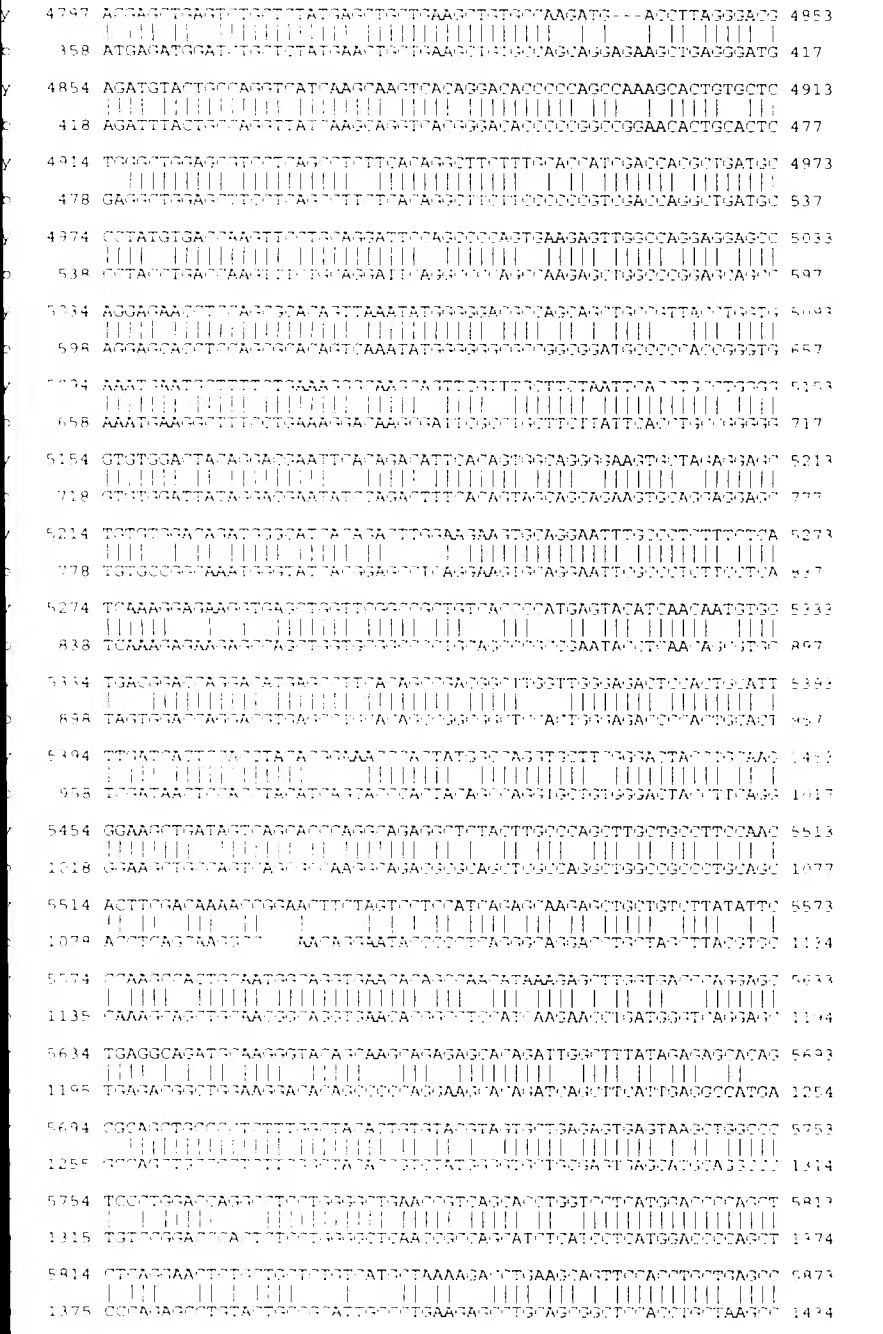


RESULT 5
ABK34864
ID ABK34864 standard; cDNA; 1960 BP.
XX
AC ABK34864;
XX
DT 18-MAY-2002 (first entry)
XX
DE Human cDNA encoding secreted protein #265.
XX
KW Human; secreted protein; gene; SP; nutritional supplement; haemophilia;
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
KW tissue regeneration; wound healing; burn; haematopoiesis;
KW myeloid cell deficiency; lymphoid cell deficiency.
XX
OS Homo sapiens.
XX
PN WO200177288 A2.
XX
PD 18-OCT 2001.
XX
PF 29-MAR-2001; 2001WO-US10224.
XX
PR 06-APR-2000; 2000US-195582P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes JB, Kenney KJ,
PI Gutshall K, Graham JR;
XX
DR WPI; 2002-175321/23.
XX
PT Five hundred and ninety two polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for
PT treating immune deficiencies and disorders such as autoimmune disorders
PT .
XX
PS Claim 1; Page 202-203; 372pp; English.
XX
CC The invention relates to 592 polynucleotides which have been derived from
CC a variety of human tissue sources and which are novel secreted
CC proteins. The polynucleotides can be used as probes for the
CC identification and isolation of full length cDNA and genomic DNA. The
CC polynucleotides and proteins can also be used as nutritional supplements.
CC The proteins are useful in the treatment of various immune deficiencies
CC and disorders such as viral infections, bacterial infections, fungal
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
CC sclerosis, autoimmune thyroiditis and diabetes), and allergic reactions
CC and conditions (e.g. asthma). They are also useful for treating
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
CC useful for tissue regeneration, for wound healing and in the treatment of
CC burns, incisions and ulcers. The proteins are also useful for regulating
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
CC Sequences ABK34861-ABK35454 represent polynucleotides of the invention.
XX
SQ Sequence 1960 BP; 433 A; 632 C; 567 G; 428 T; 0 other;



Query Match 14.5%; Score 914; DB 14; Length 1960;
Best Local Similarity 77.5%; Pred. No. 4.3e-197;
Matches 1190; Conservative 0; Mismatches 345; Indels 12; Gaps 4;



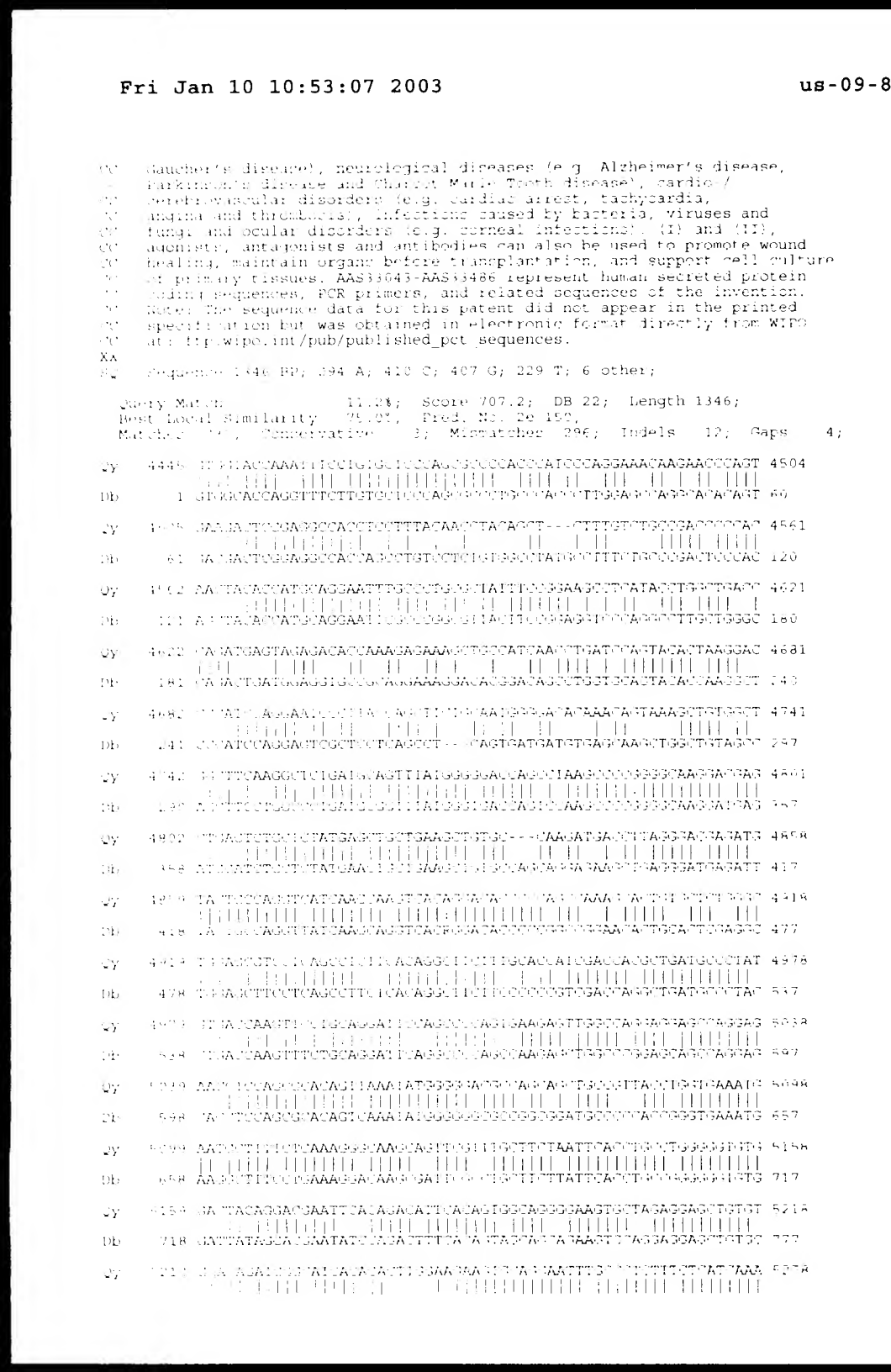


Qy 5874 CACTGCAGGAGGACGGGGCCCCCTGGCCTAGAACTCAACTATGGCTCTGTTGACAACCCCC 5933
Db 1435 CTCTGGAGGAGAAGGGGCCCCCTGGCCTGGAAGTCAACTATGGCTCAGTTGACAACCCC 1494

Qy 5934 AGACCATCTGGTTGGAGTTGCCACAGGCCCAGGAGCTGCAGCACACCATCTTCCTGC 5991
Db 1495 AGACCATCTGGTTTGAGCTGCCACAGGCCCAGGAGCTGCTATACAGCACTGTCTTCCTGA 1554

Qy 5994 TGGGCAGCATGTCCACT 6010
Db 1555 TAGACAGCAGTGCCTCT 1571

RESULT 6
AAS33274
ID AAS33274 standard; cDNA; 1346 BP.
XX
AC AAS33274;
XX
DT 04-DEC-2001 (first entry)
XX
DE DNA encoding human secreted protein, Seq ID No 233.
XX
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; anginal;
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease, neurological disease, cerebrovascular disorder;
KW thrombosis; wound healing; ss.
XX
OS Homo sapiens.
XX
PN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01347.
XX
PF 31 JAN 2000; 2000US-0176065.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451931/48.
DR P-PSDB; AAU20565.
XX
PT New nucleic acids and polypeptides, useful for diagnosing, preventing
PT or treating medical conditions -
XX
PS Claim 1; SEQ ID No 233; 753pp; English.
XX
CC The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify ligands
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example immuno/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and



```
XX
OS   Claim 9; Page 189; 229pp; English
XX
CC   AAZ56698 to AAZ56776 and AAY57877 to AAY57955 which represent human
CC   transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively.
CC   The transmembrane protein have immunospecific, antiproliferative and
CC   neuroprotective activities. The human transmembrane proteins,
CC   polynucleotides encoding them and other compositions and methods from
CC   the present invention, can be used for the diagnosis, treatment or
CC   prevention of immune, reproductive, smooth muscle, neurological,
CC   gastrointestinal, developmental and cell proliferative disorders. The
CC   HTMPN's can be used to treat or prevent disorders associated with a
CC   decreased expression or activity of HTMPN.
XX
SQ   Sequence 1476 BP; 277 A; 479 C; 435 G; 285 T; 0 other;

  Query Match             [illegible];  Score [illegible];  DB 21;  Length 1476;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  896;  Conservative    0;  Mismatches  234;  Indels   13;  Gaps    3;

Qy       2945 AA[illegible]TTGGCTAAGTTAGGGATAAATGGCGTCCACTTGCCCCTATCGA---CA 3004
Db          2 AAGGA[illegible]TATCAACGGTGCCCACTCGTCCCCGCCGATGCTG 61

Qy       3005 [illegible]TGTAGTTCCTCGACCTAAGGCTCGACCT 3064
Db         62 [illegible]AGCTTGGCTCCTCGACCCAAGGCCCCGCTA 121

Qy       3065 [illegible]AGGAAAAGCAGGGACCCCTTCGGGACTTGTTTGGC 3124
Db        122 CA[illegible]AAGGAAAAGCAGGGCCCCTTCTGGACCTGTTTGGC 181

Qy       3125 [illegible]CCACCAGCACTCCCACCG 3184
Db        182 CAGAAG[illegible]CACCTCCACCAGCGCCACCACTG 232

Qy       3185 [illegible]CATGCCTTGACAGAGCCCATG 3244
Db        233 [illegible]GTCGTTGCCTTGACACAGCCCGTG 292

Qy       3245 GA[illegible]GAAGTGGTGCTTCTCCTAT 3304
Db        293 GAGGA[illegible]TGTGCTTCT[illegible] 352

Qy       3305 [illegible]CTACCCCGGGAGAACTTC 3364
Db        353 ACCCGCACGCCCTGGAAGTTGTT[illegible]CTACCCACGGGAGAATTTC 412

Qy       3365 A[illegible]GCAGATCCTGCGGGACACCTTCACAGAG 3424
Db        413 A[illegible]GATCCTACGGGACACCTTCTCCGAG 472

Qy       3425 T[illegible]AAAATGAAAGGCCTTCTGGGAGACTTG 3484
Db        473 T[illegible]GGAAAATGAAAGACCTGCTGGGAGGCTTG 532

Qy       3485 GAGGTGAGTCTGGAGACCCTTGACATTGTTGAAGACAGCATCAAAAAATGCATCGTGGTC 3544
Db        533 GAGGT[illegible]AAGACAGCGTCAAGAAGCGCATCGTGGTG 592

Qy       3545 [illegible]CATCTTCCCAGTCTCGGGTGAGAGT 3604
Db        593 [illegible]CTGTCTCGGGCGAGAGT 652

Qy       3605 [illegible]ACCGGGGACTGCGGCTGCTGAAGGTGACC 3664
Db        653 [illegible]TGGACTGCTCAAGGTGACC 712

Qy       3665 CAAAG[illegible]CTGTTCCTACAGCTATGCTGAA 3724
Db        713 CAAGGCCCCGGCCTCCGCCCCGACCAGCTGAAGATTCTCTGCTCATACAGCTTTGCGGAG 772

Qy       3725 [illegible]CCTGGAGCTGTCCTTGAAGAATGAGCAG 3784
Db        773 GTGCTGGGTGTGGAGTGCCGGGGCGGC[illegible]GAAGAGCGAGCAG 832

Qy       3785 CTGATACTGCACACAGCCTGGGCGAGGGCCATCAAGGCCATGGTGGATCTATTTCTGAGT 3844
Db        833 CTGGTGCTGCACACAGCCCGGGCAAGGG[illegible]ATCGAGGCGCTGGTTGAGCTATTCCTGAAT 892

Qy       3845 GAACTCAGGAAGGACTCCGGCTATGTCATCGCCCTGCGCAGCTACATCAGCGATGACAAT 3904
Db        893 GAGCTTAAGAAGGACTCCGGCTATGTCATCG[illegible]TGACAACTGC 952

Qy       3905 AGCCTCCTCAGTTTCCACCGTGGGGACCTCATTAGGTTACT[illegible] 3963
Db        953 AGCCTCCTCAGCTTCCACCGTGGGGACCTCATTAAGCTG[illegible] 1012

Qy       3964 ACCAGGCTGGCAGTTCGGTTCTGCC[illegible]GATGACCTGGT 4023
Db       1013 GCCAGGCTGGCAGTTTGGCTCTGC[illegible]TGCCGACATAGT 1072

Qy       4024 GCAGCCAGCTGCTGCCCCCGACCTCTCCTTT[illegible]TGGGAAAGACAAACAGCTCGCA 4080
Db       1073 GCAGCCGGCTGCCGCTCCCGACTTTTCC[illegible]TCCAAGGA[illegible]GGCA 1129

RESULT 8
AAD07791
ID   AAD07791 standard; cDNA; 1044 BP.
XX
AC   AAD07791;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human secreted protein-encoding gene 8 cDNA clone HEONC95, SEQ ID NO:[illegible].
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   inflammation; allergy; neurological disorder; Alzheimer's disease;
KW   Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW   skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW   cardiovascular disorder; angiogenic disorder; kidney disorder;
KW   gastrointestinal disorder; pregnancy-related disorder;
KW   endocrine disorder; infection; wound healing; vulnerary;
KW   cell culture; chemotaxis; food additive; gene therapy;
KW   binding partner identification; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             616..657
FT                   /*tag=  a
FT                   /product= "Human secreted protein precursor"
FT                   /note= "CDS does not include start codon"
FT                   /partial
XX
PN   WO200132687-A1.
XX
PD   10-MAY-2001.
XX
PF   25-OCT-2000; 2000WO-US29364.
XX
PR   29-OCT-1999;   99US-0162238.
PR   30-JUN-2000; 2000US-0215134.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Duan RD,  Fiscella M,  Ebner R;
XX
DR   WPI; 2001-308740/32.
DR   P-PSDB; AAE03399.
XX
PT   Isolated nucleic acid molecule encoding a human secreted protein is
PT   used in preventing, treating or ameliorating a medical condition
```

```
XX
PS   Claim 1; Page 362; 398pp; English.
XX
CC   AAD[illegible]-AAD77794 represent cDNAs corresponding to 10 human secreted
CC   protein genes, and AAE[illegible]-AAE[illegible] represent the proteins they encode.
CC   AAE0[illegible]-AAE0341[illegible] represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing,
CC   treating or ameliorating medical conditions, e.g., by protein or gene
CC   therapy. Pathological conditions can be diagnosed by determining the
CC   amount of the new protein in a sample or by determining the presence of
CC   mutations in the new genes. Specific uses are described for each of the
CC   10 genes, based on the tissues in which they are most highly expressed,
CC   and include developing products for the diagnosis or treatment of
CC   proliferative disorders, cancer, tumours, foetal and developmental
CC   abnormalities, haematopoietic disorders, diseases of the immune system,
CC   AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC   allergies, neurological disorders (e.g., Alzheimer's disease,
CC   Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC   skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC   cardiovascular disorders, angiogenic disorders, kidney disorders,
CC   gastrointestinal disorders, pregnancy related disorders, endocrine
CC   disorders, and infections. The proteins can also be used to aid wound
CC   healing and epithelial cell proliferation, to prevent skin ageing due to
CC   sunburn, to maintain organs before transplantation, for supporting cell
CC   culture of primary tissues, to regenerate tissues, to identify their
CC   cognate ligands or binding partners, and in chemotaxis, and can be used
CC   as a food additive or preservative to modify storage properties.
CC   Antibodies specific for a protein of the invention can be used in
CC   alleviating symptoms associated with the disorders mentioned above, and
CC   in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC   immunosorbent assay (ELISA). The present sequence represents a human
CC   secreted protein encoding cDNA of the invention.
XX
SQ   Sequence 1044 BP; 217 A; 350 C; 326 G; 146 T; 5 other;

  Query Match              10.0%;  Score 628.8;  DB 22;  Length 1044;
  Best Local Similarity   76.4%;  Pred. No. 1.1e-132;
  Matches  [illegible];  Conservative    1;  Mismatches  236;  Indels    9;  Gaps    2;

Qy   [illegible] [illegible]GAGCAGTTTGCAG[illegible]GGATCCTACAGAGCAGAGGCCTGGAGGCGCCCCCTGGTG 1634
Db      7 [illegible]GAGANTTTGGNTGAATGGACCCTGTAGAGCAGAGGCCTGGAGGCGCCTCCCCGGG 66

Qy   1635 [illegible]TGTGTCACTGCATTCTGGGGATGCTTGGGTGACTTGGCTGGCTGTGACTTTG 1694
Db     67 GCTGGTCTGTGTCACTGTACTCCAGGGACGCATGGCAGGACTTGGCTGGCTGGGACTTTG 126

Qy   1695 [illegible]GAGCCTAATA[illegible]GAGGA[illegible]GAGA[illegible]GCTGGT[illegible]CAGAATAGG 1754
Db    127 [illegible]GGACCTGATCAGCCAGACTGAGGAC[illegible]GGGACCCAGCTGGCCCCGGCAGCTACC 186

Qy   1755 [illegible]ACTCCTCTTGGTTTA[illegible]TGAGAGCATC[illegible]GGTG[illegible]CAGGCTCCTT 1814
Db    187 [illegible]ACTCCTCT[illegible]GGCTCAGC[illegible]AGGCCATTCCCCTTGCCCCTGGCATTCAGGCCCCCT 246

Qy   1815 [illegible]AAATAC[illegible]AGCAG[illegible] 1874
Db    247 [illegible]CCAGGACC[illegible]CCAGGTCCAG[illegible]AACGCTGCCCAGCAGGGACCACACAG 306

Qy   1875 [illegible]CAGTAAG[illegible]TGAGAACTTGGATGGTTTCGTGGACCACCTCTTTGAACCAGCGC 1934
Db    307 G[illegible]CAGAGG[illegible]GGA[illegible]CGGACCAGATCTTCCAGCCAGTGA 366

Qy   1935 T[illegible]GGGTTTCAGTGATCTGGAACAAGGCTAGGCCTGAGCAGACGCATGAAGGGAG 1994
Db    367 TATCCTCCGGGCTCAGCGATCTGGAACAAAGCTGGGCTCTGAGCAGCCGCATGAAGGGAG 426

Qy   1995 [illegible]TCTGT[illegible]CAGGGGCTACCCCATGGTGTACCCAGGTATGGTGCAGG 2054
Db    427 [illegible]CAT[illegible]AGAGGG[illegible]TAC[illegible]CATGGTGTACCCAGGAATGATTCAGA 486

Qy   2055 [illegible]TAC[illegible]TATGATAC[illegible]CATG[illegible]CATGCCAGCCATGGGCG 2114
Db    487 TGCCTGCATACCAGCCAGGCATGGTCCCTGCAC[illegible] 546

Qy   2115 CAGTCCCAACCATGCCAGCCATGATGGTG[illegible] 2174
Db    547 [illegible]

Qy   2175 TGGACTCAAGGCAGCTGGCACTACAGCAGCAAAACTTCATCAACCAGCAGGCGATGATT[illegible] 2234
Db    601 [illegible]

Qy   2235 [illegible]CAGCAGATGACCAGCCAGGCCATGAGC[illegible] 2294
Db    661 TGGCCCAGCAGATGACAGCCCAGGCCATG[illegible] 720

Qy   2295 A[illegible]CAGCA[illegible]GGGCCCAC[illegible] 2354
Db    721 GGCAGCAGCAGGCTGGGGC[illegible]GGAGGCTGGGTG[illegible] 780

Qy   2355 [illegible]CAAGGCCAAGAAGCCTC[illegible] 2414
Db    781 [illegible] 840

Qy   2415 [illegible] 2474
Db    841 [illegible] 897

Qy   2475 [illegible]TTTTCAACAGAAATGGGACTATTTCCAGAAGAT[illegible] 2534
Db    898 GCTTCCAGCAGAAACGGAACTATTTCCAGAGGAT[illegible] 957

Qy   2535 [illegible] 2594
Db    958 [illegible]GATGAAG[illegible] 1017

Qy   2595 AGGAGGATGAGACCGCCGAGTTG 2617
Db   1018 AGGAGGAGGAGGAGGAGGAGGAGGNG 1040


RESULT 9
ABK43544
ID   ABK43544 standard; cDNA; 1768 BP.
XX
AC   ABK43544;
XX
DT   05-JUN-2002  (first entry)
XX
DE   DNA encoding novel central nervous system protein #1.4.
XX
KW   Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW   hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW   cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW   nervous system disorder; Alzheimer's disease; AIDS; [illegible] disorder;
KW   acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW   adenocarcinoma; reproductive system disorder; testicular [illegible];
KW   endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW   respiratory disorder; renal disorder; kidney failure; blood disorder;
KW   myocardial infarction; wound healing; cell proliferation; skin [illegible];
KW   food additive; food preservative; gene therapy; [illegible]; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155318-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01332.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
```

17-MAR-2000; 2000US-0140776.
18-APR-2000; 2000US-0198103.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216887.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218296.
26-JUL-2000; 2000US-0220067.
26-JUL-2000; 2000US-0220464.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225019.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225272.
14-AUG-2000; 2000US-0225441.
14-AUG-2000; 2000US-0225442.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225773.
18-AUG-2000; 2000US-0226278.
22-AUG-2000; 2000US-0226841.
22-AUG-2000; 2000US-0226868.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227366.
30-AUG-2000; 2000US-0228804.
01-SEP-2000; 2000US-0229047.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229973.
05-SEP-2000; 2000US-0229974.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231042.
08-SEP-2000; 2000US-0231043.
08-SEP-2000; 2000US-0231044.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0231980.
08-SEP-2000; 2000US-0231981.
12-SEP-2000; 2000US-0231988.
14-SEP-2000; 2000US-0232447.
14-SEP-2000; 2000US-0232448.
14-SEP-2000; 2000US-0232388.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0232403.
14-SEP-2000; 2000US-0232404.
14-SEP-2000; 2000US-0232405.
21-SEP-2000; 2000US-0234273.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235494.
27-SEP-2000; 2000US-0235934.
27-SEP-2000; 2000US-0235946.
29-SEP-2000; 2000US-0236307.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
03-OCT-2000; 2000US-0238345.

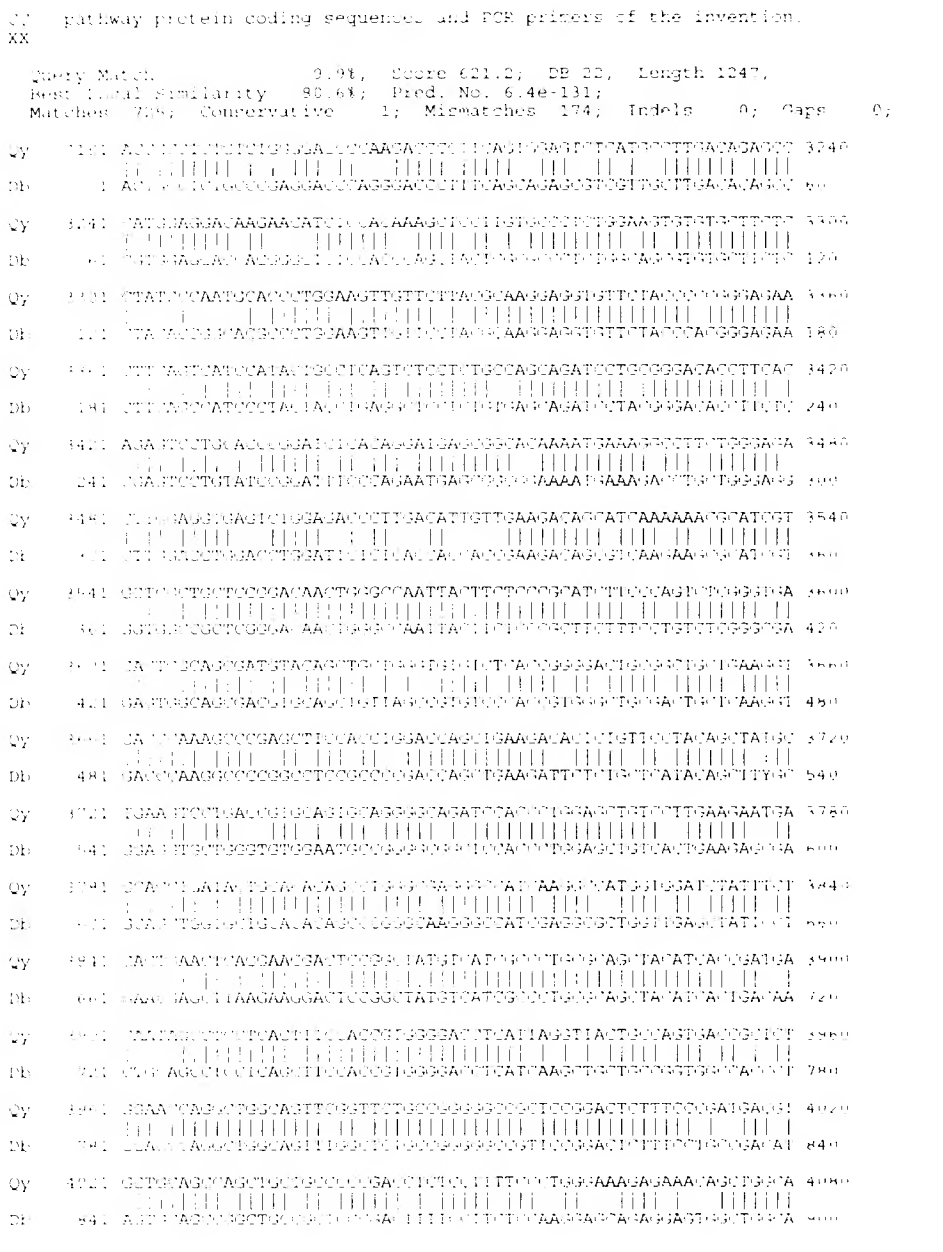
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 06-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251969.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 01-JAN-2001; 2001US-0259678.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-581633/65.
DR P-PSDB; AAU87214.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
PT preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives
XX
PS Claim 1; SEQ ID No 134; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a

```
CC   novel central nervous system protein (I) and polypeptides (III) encoded
CC   by (I), are used to treat a medical conditions and in diagnosis of a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC   amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC   e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC   e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC   adenocarcinomas and irritable bowel syndrome, reproductive system
CC   disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC   and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC   leukaemia, disorders involving neovascularisation e.g. malignancies,
CC   respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC   acute kidney failure and blood related disorders e.g. myocardial
CC   infarction. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,

  Query Match             10.0%;  Score 626.4;  DB 23;  Length 1268;
  Best Local Similarity   80.6%;  Pred. No. 4.2e-132;
  Matches  [illegible];  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy   [illegible] [illegible]AAGACTTTTAGTGGAGTCTCATGCTTTG 3232
Db         [illegible] [illegible]TGCACTGCTTCTG[illegible]GGAGGACCCAGGGATCCTTTCAGCAGAGCGTCGTTGCTTG 64

Qy   [illegible] [illegible]TGGAGGAAGTGTG 3292
Db      65 [illegible]TGGTAGCTTG 124

Qy   [illegible] [illegible]AATGCAGCCTGGAACTTGTT[illegible]AGGCAAGGAGGTGTTCTA[illegible] 3352
Db     125 [illegible]GGAAGGAGGTGTTCTAGCCA 184

Qy   [illegible] [illegible]AGAACTTCAGTCATCCATACTGCCTCAGTCTCCTCTGCCAGCAGATCCTGGGGAC 3412
Db     185 [illegible]GAGAACTTCAGCCAT[illegible]CTACTACCTGAGGCTCCTCTGTGAGCAGATCCTACGGGAC 244

Qy    3413 A[illegible]GATCTCACAGGATGAGCGGCACAAAATGAAAGGCCTT 3472
Db     245 A[illegible]GAAATGAGCGGCGGAAAATGAAAGACCTG 304

Qy    [illegible] [illegible]TGAGACAGCATCAAAAAA 3532
Db     305 CTGGGAGGCTTGGAGGTGGACCTGGATTCTCTCACCACCACGAAGACAGCGTCAAGAAG 364

Qy    [illegible] [illegible]CAATTATTTC[illegible] 3592
Db     [illegible] [illegible]GCTCCCCACAACTCCGGCCAATTACTTCTCCCGCTTCTTTCCTCTC 424

Qy    3593 [illegible]GTGACAGTGGCAGCGATGTA[illegible] 3652
Db     425 T[illegible]GCGAGAGTGGCAGCGACGTGCAGCTGTTAGCCGTGTCCCACGGTGGGCTGCGACTG 484

Qy    36[illegible] CTGAACCTCACCCAAAG[illegible]GAAGACACTCTGTT[illegible]TAC 3712
Db     485 CT[illegible]AAGGTGACCCAAGGCCCCGGCCCTCCCCCCGACCAGCTGAAGATTCTCTGCTCATAC 544

Qy    [illegible] A[illegible]GGAGCTGTCCTTG 3772
Db     545 A[illegible]TGCA[illegible]604

Qy    [illegible] AAGAATGAGCAGCTGATACTGCACACAGCCTGGGCGAGGGCCATCAAGGCCATGGTGGAT 3832
Db     [illegible] AA[illegible]664

Qy    [illegible] CTATT[illegible]CAGTA[illegible] 3892
Db     665 CTATTCCTGAATGAGCTTAA[illegible]

Qy    3893 ACTGATGACAATA[illegible] [illegible]
Db     725 ACTGACAACT[illegible]784

Qy    3953 A[illegible]GAACCAGGCTGGCAG[illegible]
Db     785 GCCACCCTGGAGCCA[illegible]TTTCCT 844

Qy    4013 [illegible]GAAAGAGAAA[illegible]
Db     845 GCCCACATACGGCA[illegible]AACAGAGAGGAT[illegible]

Qy    4073 AGCTGGCA 4080
Db     905 GGCTGGCA 912
```

```
RESULT 10
AAS27110
ID   AAS27110 standard; cDNA; 1247 BP.
XX
AC   AAS27110;
XX
DT   07 NOV 2001  (first entry)
XX
DE   DNA encoding novel signal transduction pathway protein, Seq ID 14[illegible].
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; [illegible];
KW   antiinflammatory; anti HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis [illegible]; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO US01312.
XX
PR   31-JAN-2000; 2000US 0179065.
PR   04-FEB-2000; 2000US 0180628.
PR   24-FEB-2000; 2000US 0184664.
PR   [illegible]-MAR-2000; 2000US [illegible].
PR   16-MAR-2000; 2000US 0189874.
PR   17-MAR-2000; 2000US 0190076.
PR   18-APR-2000; 2000US 0198123.
PR   19-MAY-2000; 2000US 02051[illegible].
PR   07-JUN-2000; 2000US 0209467.
PR   28-JUN-2000; 2000US 0214886.
PR   30-JUN-2000; 2000US 0215135.
PR   07-JUL-2000; 2000US 0216647.
PR   07-JUL-2000; 2000US 0216880.
PR   11-JUL-2000; 2000US 0217487.
PR   11-JUL-2000; 2000US 0217496.
PR   14-JUL-2000; 2000US 0218290.
PR   26-JUL-2000; 2000US 0220963.
PR   26-JUL-2000; 2000US 0220964.
PR   14-AUG-2000; 2000US 0224518.
PR   14-AUG-2000; 2000US 0224519.
PR   14-AUG-2000; 2000US 0225213.
PR   14-AUG-2000; 2000US 0225214.
PR   14-AUG-2000; 2000US 0225266.
PR   14-AUG-2000; 2000US 0225267.
PR   14-AUG-2000; 2000US 0225268.
```

R 14-AUG-2000; 2000US-0225270.
R 14-AUG-2000; 2000US-0225447.
R 14-AUG-2000; 2000US-0225757.
R 14-AUG-2000; 2000US-0225758.
R 14-AUG-2000; 2000US-0225929.
R 16-AUG-2000; 2000US-0226279.
R 22-AUG-2000; 2000US-0226681.
R 22-AUG-2000; 2000US-0226868.
R 22-AUG-2000; 2000US-0227182.
R 23-AUG-2000; 2000US-0227309.
R 30-AUG-2000; 2000US-0228924.
R 01-SEP-2000; 2000US-0229297.
R 01-SEP-2000; 2000US-0229343.
R 01-SEP-2000; 2000US-0229344.
R 01-SEP-2000; 2000US-0229345.
R 05-SEP-2000; 2000US-0229513.
R 06-SEP-2000; 2000US-0230457.
R 06-SEP-2000; 2000US-0230458.
R 06-SEP-2000; 2000US-0231242.
R 08-SEP-2000; 2000US-0231243.
R 08-SEP-2000; 2000US-0231244.
R 08-SEP-2000; 2000US-0231413.
R 08-SEP-2000; 2000US-0231414.
R 08-SEP-2000; 2000US-0232080.
R 08-SEP-2000; 2000US-0232081.
R 12-SEP-2000; 2000US-0231968.
R 14-SEP-2000; 2000US-0232397.
R 14-SEP-2000; 2000US-0232398.
R 14-SEP-2000; 2000US-0232399.
R 14-SEP-2000; 2000US-0232400.
R 14-SEP-2000; 2000US-0232401.
R 14-SEP-2000; 2000US-0233063.
R 14-SEP-2000; 2000US-0233064.
R 14-SEP-2000; 2000US-0233065.
R 21-SEP-2000; 2000US-0234223.
R 21-SEP-2000; 2000US-0234274.
R 25-SEP-2000; 2000US-0234997.
R 25-SEP-2000; 2000US-0234998.
R 26-SEP-2000; 2000US-0235484.
R 27-SEP-2000; 2000US-0235834.
R 27-SEP-2000; 2000US-0235836.
R 28-SEP-2000; 2000US-0236327.
R 29-SEP-2000; 2000US-0236367.
R 29-SEP-2000; 2000US-0236368.
R 29-SEP-2000; 2000US-0236369.
R 29-SEP-2000; 2000US-0236370.
R 02-OCT-2000; 2000US-0236852.
R 02-OCT-2000; 2000US-0237037.
R 02-OCT-2000; 2000US-0237038.
R 02-OCT-2000; 2000US-0237039.
R 02-OCT-2000; 2000US-0237040.
R 13-OCT-2000; 2000US-0239935.
R 13-OCT-2000; 2000US-0239937.
R 20-OCT-2000; 2000US-0241221.
R 20-OCT-2000; 2000US-0241785.
R 20-OCT-2000; 2000US-0241796.
R 20-OCT-2000; 2000US-0241797.
R 20-OCT-2000; 2000US-0241805.
R 20-OCT-2000; 2000US-0241809.
R 20-OCT-2000; 2000US-0241826.
R 01-NOV-2000; 2000US-0244617.
R 08-NOV-2000; 2000US-0246474.
R 08-NOV-2000; 2000US-0246475.
R 08-NOV-2000; 2000US-0246476.
R 08-NOV-2000; 2000US-0246477.
R 08-NOV-2000; 2000US-0246478.
R 08-NOV-2000; 2000US-0246523.
R 08-NOV-2000; 2000US-0246524.
R 08-NOV-2000; 2000US-0246525.
R 08-NOV-2000; 2000US-0246526.
R 08-NOV-2000; 2000US-0246527.

PP 08-NOV-2000; 2000US-0246528.
PP 08-NOV-2000; 2000US-0246532.
PP 08-NOV-2000; 2000US-0246609.
PP 08-NOV-2000; 2000US-0246610.
PP 08-NOV-2000; 2000US-0246611.
PP 08-NOV-2000; 2000US-0246613.
PP 17-NOV-2000; 2000US-0249207.
PP 17-NOV-2000; 2000US-0249208.
PP 17-NOV-2000; 2000US-0249209.
PP 17-NOV-2000; 2000US-0249210.
PP 17-NOV-2000; 2000US-0249211.
PP 17-NOV-2000; 2000US-0249212.
PP 17-NOV-2000; 2000US-0249213.
PP 17-NOV-2000; 2000US-0249214.
PP 17-NOV-2000; 2000US-0249215.
PP 17-NOV-2000; 2000US-0249216.
PP 17-NOV-2000; 2000US-0249217.
PP 17-NOV-2000; 2000US-0249218.
PP 17-NOV-2000; 2000US-0249244.
PP 17-NOV-2000; 2000US-0249245.
PP 17-NOV-2000; 2000US-0249264.
PP 17-NOV-2000; 2000US-0249265.
PP 17-NOV-2000; 2000US-0249297.
PP 17-NOV-2000; 2000US-0249299.
PP 17-NOV-2000; 2000US-0249300.
PP 01-DEC-2000; 2000US-0250160.
PP 01-DEC-2000; 2000US-0250391.
PP 05-DEC-2000; 2000US-0251030.
PP 05-DEC-2000; 2000US-0251988.
PP 05-DEC-2000; 2000US-0256719.
PP 06-DEC-2000; 2000US-0251479.
PP 08-DEC-2000; 2000US-0251856.
PP 08-DEC-2000; 2000US-0251868.
PP 08-DEC-2000; 2000US-0251869.
PP 08-DEC-2000; 2000US-0251989.
PP 08-DEC-2000; 2000US-0251990.
PP 11-DEC-2000; 2000US-0254097.
PP 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
DP WPI; 2001-465460/50.
DR P-PSDB; AAU17193.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 145; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T cells, to induce higher affinity antibodies, and as a means to reduce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction



RESULT 11
AR64385
ID AR64385 standard; cDNA; 1247 BP.
XX
AC AR64385;

XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #411.
XX
KW Central nervous system; CNS; autoimmune disease, rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemic angiopathy;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphasia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin ageing;
KW food additive; food preservative; gene therapy; DNA; RA.
XX
OS Homo sapiens.
XX
PN WO200155318 A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01332.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

R 08-SEP-2000; 2000US-0232080.
P 08-SEP-2000; 2000US-0232081.
R 12-SEP-2000; 2000US-0231968.
R 14-SEP-2000; 2000US-0232397.
R 14-SEP-2000; 2000US-0232398.
R 14-SEP-2000; 2000US-0232399.
R 14-SEP-2000; 2000US-0232400.
R 14-SEP-2000; 2000US-0232401.
R 14-SEP-2000; 2000US-0233063.
R 14-SEP-2000; 2000US-0233064.
R 14-SEP-2000; 2000US-0233065.
R 21-SEP-2000; 2000US-0234223.
R 21-SEP-2000; 2000US-0234274.
R 25-SEP-2000; 2000US-0234997.
R 25-SEP-2000; 2000US-0234998.
R 26-SEP-2000; 2000US-0235484.
R 27-SEP-2000; 2000US-0235834.
R 27-SEP-2000; 2000US-0235836.
R 29-SEP-2000; 2000US-0236327.
R 29-SEP-2000; 2000US-0236367.
R 29-SEP-2000; 2000US-0236368.
R 29-SEP-2000; 2000US-0236369.
R 29-SEP-2000; 2000US-0236370.
R 02-OCT-2000; 2000US-0236802.
R 02-OCT-2000; 2000US-0237037.
R 02-OCT-2000; 2000US-0237038.
R 02-OCT-2000; 2000US-0237039.
R 02-OCT-2000; 2000US-0237040.
R 13-OCT-2000; 2000US-0239935.
R 13-OCT-2000; 2000US-0239937.
R 20-OCT-2000; 2000US-0240960.
R 20-OCT-2000; 2000US-0241221.
R 20-OCT-2000; 2000US-0241785.
R 20-OCT-2000; 2000US-0241786.
R 20-OCT-2000; 2000US-0241787.
R 20-OCT-2000; 2000US-0241808.
R 20-OCT-2000; 2000US-0241809.
R 20-OCT-2000; 2000US-0241826.
R 01-NOV-2000; 2000US-0244617.
R 08-NOV-2000; 2000US-0246474.
R 08-NOV-2000; 2000US-0246475.
R 08-NOV-2000; 2000US-0246476.
R 08-NOV-2000; 2000US-0246477.
R 08-NOV-2000; 2000US-0246478.
R 08-NOV-2000; 2000US-0246523.
R 08-NOV-2000; 2000US-0246524.
R 08-NOV-2000; 2000US-0246525.
R 08-NOV-2000; 2000US-0246526.
R 08-NOV-2000; 2000US-0246527.
R 08-NOV-2000; 2000US-0246528.
R 08-NOV-2000; 2000US-0246532.
R 08-NOV-2000; 2000US-0246609.
R 08-NOV-2000; 2000US-0246610.
R 08-NOV-2000; 2000US-0246611.
R 08-NOV-2000; 2000US-0246613.
R 17-NOV-2000; 2000US-0249207.
R 17-NOV-2000; 2000US-0249208.
R 17-NOV-2000; 2000US-0249209.
R 17-NOV-2000; 2000US-0249210.
R 17-NOV-2000; 2000US-0249211.
R 17-NOV-2000; 2000US-0249212.
R 17-NOV-2000; 2000US-0249213.
R 17-NOV-2000; 2000US-0249214.
R 17-NOV-2000; 2000US-0249215.
R 17-NOV-2000; 2000US-0249216.
R 17-NOV-2000; 2000US-0249217.
R 17-NOV-2000; 2000US-0249218.
R 17-NOV-2000; 2000US-0249244.
R 17-NOV-2000; 2000US-0249245.
R 17-NOV-2000; 2000US-0249264.
R 17-NOV-2000; 2000US-0249265.
R 17-NOV-2000; 2000US-0249297.
R 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-581633/65.
DR P-PSDB; AAU87523.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
PT preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives -
XX
PS Claim 1; SEQ ID No 443; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein (II) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amylotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 9.9%; Score 621.2; DB 23; Length 1247;
Best Local Similarity 80.6%; Pred No. 6.4e-131;
Matches 725; Conservative 1; Mismatches 174; Indels 0; Gaps 0;

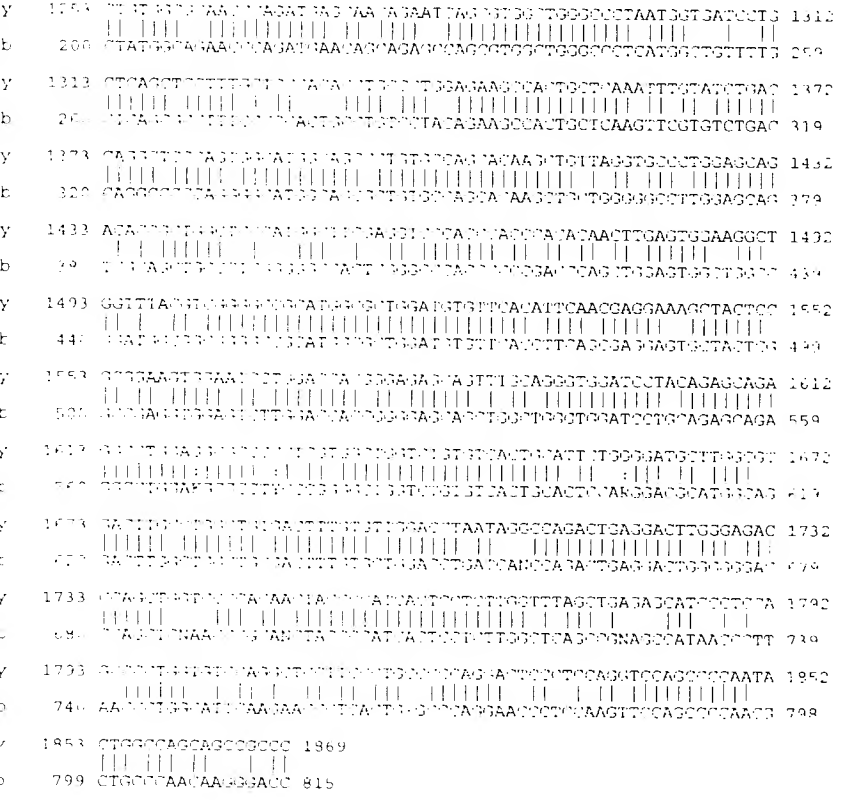
```
Qy    3181 ACGGCCTCTGTCTGGGGAGCCCAAGACCCCTTCAGTGGAGTCTCATGCCTTGACAGAGCC 3240
           || ||||||| | | ||| ||   ||||| |||||  |||  || |  ||||||| |||
Db       1 ACTGCCTCTGCCCGAGGACCCAGGGACCCTTTCAGCAGAGCGTCGTTGCTTGACACAGCC 60

Qy    3241 CATGGAGGACAAGAACATCTCCACAAAGCTCCTTGTGCCCTCTGAAAGTGTGTGCTCTCTC 3300
           | |||||||| ||    |||||||  |||| || | |||||||| || ||||||||||
Db      61 CGTGGAGGACCAGGGGTCTCCACCCAGCTACTCGCGCCCTCTGGCAGCGTGTGCTCTCTC 120

Qy    3301 CTATGCCAATGCACCCTGGAAGTTGTTCTTACGCAAGGAGGTGTTCTACCCCTGGGAGAA 3360
           |||  ||   | ||||||||||||||| ||||||||||||||||||||||| ||||||||
Db     121 CTACACGGCACGCCCTGGAAGTTGTTCCTACGCAAGGAGGTGTTCTACCCCTGGGAGAA 180

Qy    3361 CTTCAGTCATCCATACTGCCTCAGTCTCCTCTGCCAGCAGATCCTCGGGAACACCTTCAC 3420
           |||||| ||||| |||| ||| || |||||||  |||||||| ||||||||||| |
Db     181 CTTCAGCCATCCCTACTACCTGAGGCTCCTCTGTGAGCAGATCCTACGGGACACCTTCTC 240

Qy    3421 AGAGTCCTGCAGCGGATCTCACAGGATGAGGGGCACAAAATGAAAGGCTTCTGGGAGA 3480
```

```
Db        [illegible] GGATTTCCAGAATGA[illegible]GGAAAATGAAAGA[illegible] 300

Qy   [illegible] TTT[illegible]AGCTCACTCTGGA[illegible]TTGACATTGTTGAAGACAGCATCAAAAAA[illegible]CATGGT 3540

Db        [illegible] TTT[illegible]AGGTGGA[illegible]TGGA[illegible]CAGAAGACAGCGTCAAGAAGCCCATGGT 360

Qy   [illegible] GGACAACTGGG[illegible]AATTACTTCTCC[illegible]CAGTCTCGGGGGA 3600

Db        [illegible] GCTGGGACAA[illegible]TGGGCCAATTACTTCTCCGGTTCTTTCCTGTCTCGGGGCGA 420

Qy   [illegible] GA[illegible]AGGCATGTACAGCTGCT[illegible]GTGTCTCA[illegible]ACTG[illegible]GTGCTGAAGG[illegible] 3660

Db        [illegible] GA[illegible]GAGTGCA[illegible]TGTTAGCCGTGTCCCACCGTGGGCTGCGACTG[illegible]AAGG[illegible] 480

Qy   [illegible] GA[illegible]AAAG[illegible]CAGCTT[illegible]GAAGACA[illegible]TGTT[illegible]TA[illegible]TAT[illegible] 3720

Db        [illegible] GA[illegible]AAGG[illegible]TGGCC[illegible]GACCAG[illegible]TGAAGATTCTCTGCTCATACAGCTTYGC 540

Qy   [illegible] TGAAGTCCTGACCGTGCAGTGCAG[illegible]CAGAT[illegible]AGCTGTCCTTGAAGAATGA 3780

Db        [illegible] GGA[illegible]TGGGTGTGGAAT[illegible]CCTCCACC[illegible]AGCTGTCACTGAAGAGCGA 600

Qy   [illegible] CCAG[illegible]TGATACTGCACACAGCCTGGGCGAGGG[illegible]ATCAAGG[illegible]CAT[illegible]GGATCTATTTCT 3840

Db        [illegible] GCAG[illegible]TGGTGCTGCACACAGCCCGGGCAAGGG[illegible]CATCGAGGCGCTGGTTGAGCTATTCCT 660

Qy   [illegible] GACT[illegible]AACT[illegible]AGGAAGGACTCCGGCTATGTCATCGCCCTGC[illegible]AGCTACAT[illegible]GATGA 3900

Db        [illegible] GAA[illegible]AGCTTAAGAAGGACTCCGGCTATGTCATCGCCCTGCGCAGCTACATCACTGACAA 720

Qy   [illegible] CAATA[illegible]CTCAGTTT[illegible]CACCGTGGGGACCTCATTAGGTTACTG[illegible]AGTGACCGCTCT 3960

Db        [illegible] CT[illegible]CAGTTT[illegible]GAAG[illegible]AGGTGGCCACCCT 780

Qy   [illegible] [illegible]AGGCTGGCAGTT[illegible]GGTTCTG[illegible]GGT[illegible]GGAT[illegible]TTT[illegible]ATGA[illegible] 4020

Db        [illegible] GGA[illegible]CAGGCTGGCAGTT[illegible]GGTT[illegible]GGAT[illegible]TTT[illegible]GGACAT 840

Qy   [illegible] GCT[illegible]ACCCA[illegible]CTGCTGGC[illegible]CCGACCTCTCC[illegible]TCCCTGGGAAAGAGAAACAGCTGGCA 4080

Db        [illegible] AG[illegible]AGCTGG[illegible]GCGGT[illegible]GGACTTTTCCTTCTCCAAGGAGCAGAGGAGTGGCTGGCA 900
```

RESULT 17
AA[illegible]
ID AA[illegible] standard; cDNA; 826 BP.
XX
AC AAD07777;
XX
DT [illegible] (first entry)
XX
DE Human secreted protein encoding gene 8 cDNA clone HFCN[illegible], SEQ ID NO:18.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnerary;
KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 171..653
FT /*tag= a
FT /product= "Human secreted protein precursor"
FT /transl_except= (pos:567..569, aa:Xaa)
FT /transl_except= (pos:6[illegible], aa:Xaa)
FT /note= "Xaa equals any of the naturally occurring
FT L-amino acids"
FT sig_peptide 171..293
FT /*tag= b
FT mat_peptide 294..650
FT /*tag= c
FT /product= "Mature human secreted protein"
XX
PN WO200132687-A1.
XX
PD 10-MAY-2001.
XX
PF [illegible]-OCT-2000; 2000WO-US29364.
XX
PR 29-OCT-1999; 99US-0162238.
PR 30-JUN-2000; 2000US-0215134.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis GA, Duan RD, Fiscella M, Ebner R;
XX
DR WPI; 2001-308740/32.
DR P-PSDB; AAE03385.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; Page 351; 398pp; English.
XX
CC AAD07770-AAD07794 represent cDNAs corresponding to 1 human secreted
CC protein genes, and AAE03378-AAE03402 represent the proteins they encode.
CC AAE03403-AAE03415 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 10 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis, inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin ageing due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein encoding cDNA of the invention.
XX
SQ Sequence 826 BP; [illegible] A; [illegible] C; [illegible] G; [illegible] T; [illegible] other;

Query Match 7.5%; Score 470.[illegible]; DB [illegible]; Length 826;
Best Local Similarity 77.7%; Pred. No. [illegible];
Matches 572; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

```
Qy   1133 CTCCTGGGGGAAGGATCTCTGCAGTCCTGG[illegible]AAGA[illegible] [illegible]

Db     80 [illegible]GGA[illegible]GGAT[illegible] [illegible]

Qy   1193 CAGCAGGCCCAGCGACGGCCGGGACTCCGAGATGA[illegible] [illegible]

Db    140 [illegible]AGGGCCA[illegible]TGCCACCCTCCGAATGA[illegible] [illegible]
```



RESULT 13
AAS27687/c
ID AAS27687 standard; DNA; 24757 BP.
XX
AC AAS27687;
XX
DT 07-NOV-2001 (first entry)
XX
DE DNA encoding novel signal transduction pathway protein, Seq ID 1347.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive, tumour;
KW antiinflammatory; anti HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing, endocrine, Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 15-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 08-JUN-2000; 2000US-0210886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227192.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229087.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

```
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   19-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465460/50.
XX
PT   Novel polypeptides useful for diagnosing, treating, preventing and/or
PT   prognosing disorders related to the proteins, including cancers, immune
PT   disorders and neuronal disorders
XX
PS   Claim 1; SEQ ID No 1347; 880pp; English.
XX
CC   The invention relates to novel isolated polypeptides (I), and
CC   polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC   diagnosing, preventing and treating diseases including immune system
CC   disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC   disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC   transplant rejections and graft versus host disease, infectious diseases
CC   (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC   other blood-related disorders (sickle cell anaemia), myeloproliferative
CC   disorders, primary haematopoietic disorders, hyperproliferative
CC   disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC   disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC   abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC   disorders (e.g. glomerulonephritis), cardiovascular disorders
CC   (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC   wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC   Addison's disease), reproductive system disorders, gastrointestinal
CC   disorder (inflammatory disorders), liver disorders (cirrhosis),
CC   as stimulators of B-cell responsiveness to pathogens, activators of
CC   T cells, to induce higher affinity antibodies, and as a means to inhibit
CC   tumour proliferation in pathologies e.g. acquired immune deficiency
CC   syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC   pathway protein coding sequences and PCR primers of the invention.
XX
SQ   Sequence 24757 BP; 5183 A; 7437 C; 7328 G; 4809 T; 0 other;

  Query Match             2.6%;  Score 166.2;  DB 22;  Length 24757;
  Best Local Similarity   80.2%;  Pred. No. 4.2e-17;
  Matches  195;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy    1119 GGGTGATCCTGCGGCT[illegible]

Db   17652 GGCAGATGCTGCGGCT[illegible]

Qy    1179 GCACGTTCCTCGTGCAGCAGGC[illegible]

Db   17592 GCGCATACCTGGTGCGGCAGGGGCAG[illegible]

Qy    1239 AGCTGGTGGCCCAGCTGTGGC[illegible]

Db   17532 AGCTGGTGGCCCAGCTATGGCAGAA[illegible]

Qy    1299 [illegible]

Db   17472 TCATGGCTGTTTTGCTCAGCGCCTT[illegible] 17413

Qy    1359 AAT 1361

Db   17412 AGT 17410


RESULT 14
AAS33481
ID   AAS33481 standard; DNA; 24757 BP.
XX
AC   AAS33481;
XX
DT   04-DEC-2001  (first entry)
XX
DE   DNA encoding human secreted protein, Seq ID No 764.
XX
KW   Immunomodulatory; human immunodeficiency virus; HIV; anaemia; asthma;
KW   rheumatoid arthritis; antiarteriosclerotic; antiasthmatic; vasotropic;
KW   cerebroprotective, thrombolytic, antimicrobial; ophthalmological;
KW   cytostatic, Alzheimer's disease, Parkinson's disease; human; cancer,
KW   multiple sclerosis, cancer; hyperproliferative disorder; infection;
KW   Gaucher's disease; neurological disease; cerebrovascular disorder;
KW   thrombosis; wound healing; ds.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200155326-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01347.
XX
PR   31-JAN-2000; 2000US-0179065.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451931/48.
XX
PT   New nucleic acids and polypeptides, useful for diagnosing, preventing
PT   or treating medical conditions
XX
PS   Disclosure; SEQ ID No 764; 753pp; English.
XX
CC   The invention relates to novel isolated nucleic acid molecules (I)
CC   encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC   treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC   goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC   the prevention, treatment and diagnosis of diseases associated with
CC   inappropriate expression of secreted proteins. (I) and complementary
CC   sequences may also be used as DNA probes in diagnostic assays (e.g.
CC   polymerase chain reactions (PCR)) to detect and quantitate the presence
CC   of similar nucleic acid sequences in samples, and so which patients may
CC   be in need of restorative therapy. (II) may also be used as antigens in
CC   the production of antibodies and in assays to identify modulators
CC   (agonists and antagonists) of the expression and activity of the secreted
CC   proteins. The anti-(II) antibodies and antagonists may also be used to
CC   down regulate expression and activity of (II). The anti-(II) antibodies
CC   may also be used as diagnostic agents for detecting the presence of (II)
CC   in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC   disorders include for example: immune/autoimmune diseases (e.g. HIV
CC   (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC   and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC   melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC   Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC   Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC   cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
CC   angina and thrombosis), infections caused by bacteria, viruses and
CC   fungi and ocular disorders (e.g. corneal infections). (I) and (II),
CC   agonists, antagonists and antibodies can also be used to promote wound
CC   healing, maintain organs before transplantation, and support cell culture
CC   of primary tissues. AAS33043-AAS33486 represent human secreted protein
CC   coding sequences, PCR primers, and related sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO
CC   at: ftp wipo int/pub/published_pct_sequences.
XX
SQ   Sequence 24757 BP; 4809 A; 7229 C; 7437 G; 5183 T; 0 other;

 Query Match             2.6%;  Score 166.2;  DB 22;  Length 24757;
 Best Local Similarity   80.2%;  Pred. No. 4.2e-27;
 Matches  195;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy    1119 GGGTGATGCTGCTGGCTGCTGGGGCAAGGAATCTGTGCAGTCCTGGCAAGAGCAGACCATGG 1178
           ||  |||||||||||||||||||||| ||||| ||| |||||||||    ||||| |||||
Db    7195 TGTAGATGCTGTGGTCTTTGGGGACTGAATGCCTGTAGTCCTGGCAGAGGCAGATCATGG 7166

Qy    1179 GCACGTTCCTCGTGCAGCAGGGCCAGGAGGGCGGGAGTCCGAGATGAGCTCTTCAGCC 1238
           ||  | | ||| |||| |||||  ||| | ||||| || || ||  |||||||||| ||||
Db    7166 GCGCATACCTGATCGGGAAGGGGCAGGCCTGCGGAATGAGCTCTTTAGCC 7225

Qy    1239 AGCTGGTGGGCCAGCTGTGGTCAAGCCAGATGAGAACAGAATCAGCGTGGCTGGGCCC 1298
           ||||||||||||||| |||| |||||||||| || ||||  ||||||||||||||||
Db    7226 AGCTGGTGGCCAGCTATGGCAGAACCAGATGAACAGCAGAGCCAGCGTGGCTGGGCCC 7285

Qy    1299 TAATGGTGATCCTGCTCAGCTGCTTTGCTCCAACCTGCGCTGGAGAAGCCACTGCTCA 1358
           |  |||||   | |||||||| ||||| | ||    |||| |||  ||||||||||||||
Db    7286 TCATGGCTGTTTTGCTCAGCGCCTTTCCCCCACTGCCTGTCCTACAGAAGCCACTGCTCA 7345

Qy    1359 AAT 1361
           | ||
Db    7346 AGT 7348


RESULT 15
AAS33480
ID   AAS33480 standard; DNA; 9703 BP.
XX
AC   AAS33480;
XX
DT   04-DEC-2001  (first entry)
XX
DE   DNA encoding human secreted protein, Seq ID No 763.
XX
KW   Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW   rheumatoid arthritis, antiarteriosclerotic; cardiant; vascular;
KW   cerebroprotective; thrombolytic, antimicrobial, ophthalmological;
KW   cytostatic, Alzheimer's disease, Parkinson's disease, human, cancer;
KW   multiple sclerosis; cancer; hyperproliferative disorder, infection;
KW   Gaucher's disease; neurological disease, cerebrovascular disorder;
KW   thrombosis; wound healing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155326-A2.
XX
PD   02 AUG-2001.
XX
PF   17-JAN 2001; 2001WO-US01347.
XX
PR   31 JAN 2000; 2000US-0179065.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI, 2001 451931/48.
XX
PT   New nucleic acids and polypeptides, useful for diagnosing, preventing
PT   or treating medical conditions -
XX
PS   Disclosure, SEQ ID No 763, 753pp; English.
XX
CC   The invention relates to novel isolated nucleic acid molecules (I)
CC   encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC   treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC   goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC   the prevention, treatment and diagnosis of diseases associated with
CC   inappropriate expression of secreted proteins. (I) and complementary
CC   sequences may also be used as DNA probes in diagnostic assays (e.g.
CC   polymerase chain reactions (PCR)) to detect and quantitate the presence
CC   of similar nucleic acid sequences in samples, and so which patients may
CC   be in need of restorative therapy. (II) may also be used as antigens in
CC   the production of antibodies and in assays to identify modulators
CC   (agonists and antagonists) of the expression and activity of the secreted
CC   proteins. The anti-(II) antibodies and antagonists may also be used to
CC   down regulate expression and activity of (II). The anti-(II) antibodies
CC   may also be used as diagnostic agents for detecting the presence of (II)
CC   in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC   disorders include for example: immune/autoimmune diseases (e.g. HIV
CC   (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC   and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC   melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC   Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC   Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC   cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
CC   angina and thrombosis), infections caused by bacteria, viruses and
CC   fungi and ocular disorders (e.g. corneal infections). (I) and (II),
CC   agonists, antagonists and antibodies can also be used to promote wound
```

XX Devising, maintain organs before transplantation, and support cell culture
XX of primary tissues. AAS33043-AAS33486 represent human secreted protein
CC coding sequences, PCR primers, and related sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 9703 BP; 1950 A; 2842 C; 3079 G; 1832 T; 0 other;

Query Match 1.6%; Score 97.6; DB 22; Length 9703;
Best Local Similarity 77.6%; Pred. No. 1.1e-11;
Matches 119; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 5809 CAGCTCTCAGGAACTCTGCTGCTCTGTCATGCTAAAAGACCTGAAGCAGTTCCACCTGCT 5868
Db 9147 CTGGCCCCAGAGCCTGTACTGCCGCATTGCCCTCAAGAGCCTGCAGGGCTCCTACCTGCT 9206

Qy 5869 GTCAGCCCAGGAGGACTGGGCCCCTTGGCCTAGAACTCAACTATGGCTCTGTTGACAA 5928
Db 9207 GTCTTCTTGGAGGAGAAGGGGCCCCTGGCCTGGAAGTCAACTATGGCTCAGCTGACAA 9266

Qy 5929 CTCAACCATCTGGTTGGAGTTGCCACAGG 5960
Db 9267 CTCATACCATCTGGTTTGAGCTGCCACAGG 9298

Search completed: January 9, 2003, 07:12:35
Job time : 147.941 secs